MATH1030 Dimension relation on sum and intersection.

1. Recall the definition for the respective notions of *intersection* and *sum* for *subspaces of* $\mathbb{R}^n$:

   Let Y, Z be subspaces of $\mathbb{R}^n$.

   (a) *The intersection of Y, Z, denoted by $V \cap W$, is the subspace of $\mathbb{R}^n$ defined to by*

   $$Y \cap Z = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \in Y \textit{ and } \mathbf{x} \in Z\}.$$

   (b) *The sum of Y, Z, denoted by $Y + Z$, is the subspace of $\mathbb{R}^n$ defined by*

   $$Y + Z = \left\{ \mathbf{x} \in \mathbb{R}^n : \begin{array}{l} \textit{There exist some } y \in Y,\ z \in Z \\ \textit{such that } x = y + z \end{array} \right\}.$$

   Also recall the result below, which is Theorem (G) from the handout *More on minimal spanning sets*:

   Let W be a non-zero subspace of $\mathbb{R}^n$, and $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ be vectors in W.
   Further suppose that $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ are linearly independent.
   Then, there is some basis for W, which is constituted of at most n vectors, amongst them being the vectors $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$.

   With the help of this result, we are going to establish an numerical equality relating the respective dimensions of any two subspaces of $\mathbb{R}^n$ and the respective dimensions of their intersection and their sum.

2. **Theorem (L). (Dimension Theorem relating the sum and intersection of subspaces of $\mathbb{R}^n$.)**

   Suppose Y, Z are subspaces of $\mathbb{R}^n$.

   Then $\dim(Y + Z) + \dim(Y \cap Z) = \dim(Y) + \dim(Z)$.

   Remark. What is nice about Theorem (L) is that it provides a relation from which we can deduce the dimension of a certain subspace of $\mathbb{R}^n$ (or obtain some constraints on its dimension), without having to go into the trouble of immediately finding a basis for it, as long as we are provided enough information on some other subspaces of $\mathbb{R}^n$.

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

   Suppose Y, Z are subspaces of $\mathbb{R}^n$. Write $\dim(Y) = m$, and $\dim(Z) = n$.

   Note that $Y \cap Z$, $Y + Z$ are subspaces of $\mathbb{R}^n$. Write $\dim(Y \cap Z) = p$, and $\dim(Y + Z) = q$.

   Pick some basis for $Y \cap Z$, say, $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_p$.

   By Theorem (G), there is some basis for Y, which is constituted by vectors in Y, amongst them being $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_p$. Denote the other vectors in this basis for Y by $\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_k$.

   By construction, we have $p + k = m$. Also, by construction, none of $\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_k$ belongs to Z. Justification:

   - Suppose it were true that $\mathbf{s}_1$ belonged to Z. Then $\mathbf{s}_1 \in Y$ and $\mathbf{s}_1 \in Z$. Therefore $\mathbf{s}_1 \in Y \cap Z$.
     Since $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_p$ constitute a basis for $Y \cap Z$, it would happen that $\mathbf{s}_1$ was a linear combination of $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_p$.
     Then $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_p, \mathbf{s}_1$ would be linearly dependent.
     However, because $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_p, \mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_k$ constitute a basis for Y, they are linearly independent. Then $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_p, \mathbf{s}_1$ are linearly dependent. Contradiction arises.
     Hence in the first place $\mathbf{s}_1$ does not belong to Z. Similarly, none of $\mathbf{s}_2, \cdots, \mathbf{s}_k$ belong to Z.

   Similarly, there is some basis for Z, which is constituted by vectors in Z, amongst them being $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_p$. Denote the other vectors in this basis for Y by $\mathbf{t}_1, \mathbf{t}_2, \cdots, \mathbf{t}_\ell$.

   By construction, we have $p + \ell = n$. Moreover, by construction, none of $\mathbf{t}_1, \mathbf{t}_2, \cdots, \mathbf{t}_\ell$ belong to Y.

   We verify that $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_p, \mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_k, \mathbf{t}_1, \mathbf{t}_2, \cdots, \mathbf{t}_\ell$ constitute a basis for $V + W$:

   - Pick any $\mathbf{u} \in Y + Z$. By definition, there exist some $\mathbf{v} \in Y$, $\mathbf{w} \in Z$ such that $\mathbf{u} = \mathbf{v} + \mathbf{w}$.
     Since $\mathbf{v} \in Y$, there exist some $\alpha_1, \alpha_2, \cdots, \alpha_p, \beta_1, \beta_2, \cdots, \beta_k$ such that

     $$\mathbf{v} = \alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \cdots + \alpha_p\mathbf{x}_p + \beta_1\mathbf{s}_1 + \beta_2\mathbf{s}_2 + \cdots + \beta_k\mathbf{s}_k.$$

     Since $\mathbf{w} \in Z$, there exist some $\gamma_1, \gamma_2, \cdots, \gamma_p, \delta_1, \delta_2, \cdots, \delta_\ell$ such that

     $$\mathbf{w} = \gamma_1\mathbf{x}_1 + \gamma_2\mathbf{x}_2 + \cdots + \gamma_p\mathbf{x}_p + \delta_1\mathbf{t}_1 + \delta_2\mathbf{t}_2 + \cdots + \delta_\ell\mathbf{t}_\ell.$$

     Then

     $$\begin{aligned} \mathbf{u} &= \mathbf{v} + \mathbf{w} \\ &= (\alpha_1 + \gamma_1)\mathbf{x}_1 + (\alpha_2 + \gamma_2)\mathbf{x}_2 + \cdots + (\alpha_p + \gamma_p)\mathbf{x}_p + \beta_2\mathbf{s}_2 + \cdots + \beta_k\mathbf{s}_k + \delta_1\mathbf{t}_1 + \delta_2\mathbf{t}_2 + \cdots + \delta_\ell\mathbf{t}_\ell \end{aligned}$$

- Pick any $\alpha_1, \alpha_2, \cdots, \alpha_p, \beta_1, \beta_2, \cdots, \beta_k, \gamma_1, \gamma_2, \cdots, \gamma_\ell \in \mathbb{R}$.
  Suppose $\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \cdots + \alpha_p\mathbf{x}_p + \beta_1\mathbf{s}_1 + \beta_2\mathbf{s}_2 + \cdots + \beta_k\mathbf{s}_k + \gamma_1\mathbf{t}_1 + \gamma_2\mathbf{t}_2 + \cdots + \gamma_\ell\mathbf{t}_\ell = \mathbf{0}$.
  Write $\mathbf{x} = \alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \cdots + \alpha_p\mathbf{x}_p$, $\mathbf{s} = \beta_1\mathbf{s}_1 + \beta_2\mathbf{s}_2 + \cdots + \beta_k\mathbf{s}_k$, $\mathbf{t} = \gamma_1\mathbf{t}_1 + \gamma_2\mathbf{t}_2 + \cdots + \gamma_\ell\mathbf{t}_\ell$.
  By assumption, we have $\mathbf{x} + \mathbf{s} + \mathbf{t} = \mathbf{0}$.
  We have $\mathbf{s} = -\mathbf{x} - \mathbf{t}$. Since $\mathbf{s} \in Y$ and $-\mathbf{x} - \mathbf{t} \in Z$, we have $\mathbf{s} \in Y \cap Z$.
  Since $\mathbf{s} \in Y \cap Z$, there exist some $\delta_1, \delta_2, \cdots, \delta_p \in \mathbb{R}$ such that $\mathbf{s} = \delta_1\mathbf{x}_1 + \delta_2\mathbf{x}_2 + \cdots + \delta_p\mathbf{x}_p$.
  Then $\delta_1\mathbf{x}_1 + \delta_2\mathbf{x}_2 + \cdots + \delta_p\mathbf{x}_p + (-\beta_1)\mathbf{s}_1 + (-\beta_2)\mathbf{s}_2 + \cdots + (-\beta_k)\mathbf{s}_k = \mathbf{s} - \mathbf{s} = \mathbf{0}$.
  Since $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_p, \mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_k$ constitute a basis for Y, we have $\delta_1 = \delta_2 = \cdots = \delta_p = 0$ and $-\beta_1 = -\beta_2 = \cdots = -\beta_k = 0$. Then $\beta_1 = \beta_2 = \cdots = \beta_k = 0$, and $\mathbf{s} = \mathbf{0}$.
  Now we have $\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \cdots + \alpha_p\mathbf{x}_p + \gamma_1\mathbf{t}_1 + \gamma_2\mathbf{t}_2 + \cdots + \gamma_\ell\mathbf{t}_\ell = \mathbf{x} + \mathbf{t} = \mathbf{0}$.
  Since $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_p, \mathbf{t}_1, \mathbf{t}_2, \cdots, \mathbf{t}_\ell$ constitute a basis for Z, we have $\alpha_1 = \alpha_2 = \cdots = \alpha_p = 0$ and $\gamma_1 = \gamma_2 = \cdots = \gamma_\ell = 0$.

Therefore $\dim(Y + Z) = p + k + \ell$.

Hence $\dim(Y + Z) + \dim(Y \cap Z) = (p + k + \ell) + p = (p + k) + (p + \ell) = \dim(Y) + \dim(Z)$.

4. **Illustrations of Theorem (L).**

(a) Let $\mathbf{u}_1 = \mathbf{e}_1^{(3)}, \mathbf{u}_2 = \mathbf{e}_2^{(3)}$, and $\mathbf{v}_1 = \mathbf{e}_2^{(3)}, \mathbf{v}_2 = \mathbf{e}_3^{(3)}$.
Suppose $Y = \mathsf{Span}\ (\{\mathbf{u}_1, \mathbf{u}_2\})$ and $Z = \mathsf{Span}\ (\{\mathbf{v}_1, \mathbf{v}_2\})$.
Then $Y + Z = \mathsf{Span}\ (\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{v}_1, \mathbf{v}_2\}) = \mathsf{Span}\ (\{\mathbf{e}_1^{(3)}, \mathbf{e}_2^{(3)}, \mathbf{e}_3^{(3)}\}) = \mathbb{R}^3$. (Why?)
Note that $\dim(Y) = 2$, $\dim(Z) = 2$ and $\dim(Y + Z) = 3$. (Why?)
Without finding a basis for $Y \cap Z$, we see that $\dim(Y \cap Z) = \dim(Y) + \dim(Z) - \dim(Y + Z) = 1$.
It happens that $Y \cap Z = \mathsf{Span}\ (\{\mathbf{e}_2^{(3)}\})$. Justification.

- Note that the non-zero vector $\mathbf{e}_2^{(3)}$ is one linearly independent vector in $Y \cap Z$. Since $\dim(Y \cap Z) = 1$, a basis for $Y \cap Z$ is constituted by $\mathbf{e}_2^{(3)}$.

(b) Let $\mathbf{u}_1 = \mathbf{e}_1^{(4)}, \mathbf{u}_2 = \mathbf{e}_2^{(4)}$, and $\mathbf{v}_1 = \mathbf{e}_3^{(4)}, \mathbf{v}_2 = \mathbf{e}_4^{(4)}$.
Suppose $Y = \mathsf{Span}\ (\{\mathbf{u}_1, \mathbf{u}_2\})$ and $Z = \mathsf{Span}\ (\{\mathbf{v}_1, \mathbf{v}_2\})$.
Then $Y + Z = \mathsf{Span}\ (\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{v}_1, \mathbf{v}_2\}) = \mathsf{Span}\ (\{\mathbf{e}_1^{(4)}, \mathbf{e}_2^{(4)}, \mathbf{e}_3^{(4)}, \mathbf{e}_4^{(4)}\}) = \mathbb{R}^4$. (Why?)
Note that $\dim(Y) = 2$, $\dim(Z) = 2$ and $\dim(Y + Z) = 4$. (Why?)
Without finding a basis for $Y \cap Z$, we see that $\dim(Y \cap Z) = \dim(Y) + \dim(Z) - \dim(Y + Z) = 0$.
It happens that $Y \cap Z = \{\mathbf{0}_4\}$, and its basis is the empty set.

(c) Let $\mathbf{s}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{s}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{t}_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \mathbf{t}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.
Define $V = \mathsf{Span}\ (\{\mathbf{s}_1, \mathbf{s}_2\})$, $W = \mathsf{Span}\ (\{\mathbf{t}_1, \mathbf{t}_2\})$.
Note that $\dim(V) = 2$ and $\dim(W) = 2$.
By definition, $V + W = \mathsf{Span}\ (\{\mathbf{s}_1, \mathbf{s}_2, \mathbf{t}_1, \mathbf{t}_2\})$.
$\mathbf{s}_1, \mathbf{s}_2, \mathbf{t}_1$ are linearly independent. (Why? How?)
Then $V + W = \mathbb{R}^3$ and $\dim(V + W) = 3$. (Why?)
Therefore $\dim(V \cap W) = 1$.

(d) Let $\mathbf{s}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{s}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{s}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \mathbf{t}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \mathbf{t}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \end{bmatrix}, \mathbf{t}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}$.
Define $V = \mathsf{Span}\ (\{\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3\})$, $W = \mathsf{Span}\ (\{\mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_3\})$.
Note that $\dim(V) = 3$ and $\dim(W) = 3$ (Why?)
By definition, $V + W = \mathsf{Span}\ (\{\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3, \mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_3\})$.
$\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3, \mathbf{t}_2$ are linearly independent. (Why? How?)
Then $V + W = \mathbb{R}^4$ and $\dim(V + W) = 4$. (Why?)
Therefore $\dim(V \cap W) = 2$.

(e) Let $\mathbf{s}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$, $\mathbf{s}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{t}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}$, $\mathbf{t}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}$.

Define $V = \mathsf{Span}\ (\{\mathbf{s}_1, \mathbf{s}_2\})$, $W = \mathsf{Span}\ (\{\mathbf{t}_1, \mathbf{t}_2\})$.

Note that $\dim(V) = 3$ and $\dim(W) = 3$. (Why?)

By definition, $V + W = \mathsf{Span}\ (\{\mathbf{s}_1, \mathbf{s}_2, \mathbf{t}_1, \mathbf{t}_2\})$.

We determine the dimension of $V + W$ by finding a basis for it:

- Write $U = [\ \mathbf{s}_1 \mid \mathbf{s}_2 \mid \mathbf{t}_1 \mid \mathbf{t}_2\ ]$.
  Apply row operations on U to find the reduced row-echelon form U' which is row-equivalent to U:

$$U = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 1 & 0 & -1 \end{bmatrix} \longrightarrow \cdots\cdots \longrightarrow U' = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

  The pivot columns of U' are the first, second and third columns.
  Then a basis for $V + W$ is constituted by $\mathbf{s}_1, \mathbf{s}_2, \mathbf{t}_1$.
  Therefore $\dim(V + W) = 3$.

It follows that $\dim(V \cap W) = 1$.

(f) Let $\mathbf{s}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{s}_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}$, $\mathbf{t}_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}$, $\mathbf{t}_2 = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}$.

Define $V = \mathsf{Span}\ (\{\mathbf{s}_1, \mathbf{s}_2\})$, $W = \mathsf{Span}\ (\{\mathbf{t}_1, \mathbf{t}_2\})$.

Note that $\dim(V) = 2$ and $\dim(W) = 2$. (Why?)

By definition, $V + W = \mathsf{Span}\ (\{\mathbf{s}_1, \mathbf{s}_2, \mathbf{t}_1, \mathbf{t}_2\})$.

It happens that $\mathbf{s}_1, \mathbf{s}_2, \mathbf{t}_1, \mathbf{t}_2$ are linearly independent.

Then $V + W = \mathbb{R}^4$, and $\dim(V + W) = 4$.

Therefore $\dim(V \cap W) = 0$ and $V \cap W = \{\mathbf{0}_4\}$.

5. We now also recall Theorem (K) (which we call the Rank-nullity Formula) from the handout *Rank-nullity Formula*:

   Let A be an $(p \times q)$-matrix.
   Denote by A' the reduced row-echelon form which is row equivalent to A, and suppose the rank of A' is $r(A)$.
   Then the statements below hold:
   (a) $r(A) = r_{col}(A) = r_{row}(A)$.
   (b) $n(A) + r(A) = q$.
   (c) $r(A^t) = r(A)$, *and* $n(A^t) + r(A) = p$.

   We shall freely apply this result in the various examples below.

6. **Further illustrations of Theorem (L).**

(a) Let $B = \begin{bmatrix} 1 & 2 & 2 & 4 \\ 1 & 3 & 3 & 5 \end{bmatrix}$, and $C = [\ 2 \quad 6 \quad 5 \quad 6\ ]$.

Define $V = \mathcal{N}(B)$ and $W = \mathcal{N}(C)$.

We have $\dim(V) + r(B) = 4$. Note that $r(B) = r_{row}(B) = 2$. Then $\dim(V) = 2$.

We have $\dim(W) + r(C) = 4$. Note that $r(C) = r_{row}(C) = 1$. Then $\dim(W) = 3$.

Define $A = \left[\dfrac{B}{C}\right]$. Note that $\mathcal{N}(A) = \mathcal{N}(B) \cap \mathcal{N}(C) = V \cap W$.

Then $\dim(V \cap W) + r(A) = 4$. We determine the value of $r(A)$:

- We obtain the reduced row-echelon form A' which is row-equivalent to A by applying a sequence of row operations to A:

$$A \longrightarrow \cdots \longrightarrow \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & 4 \end{bmatrix} = A'$$

  We see that $r(A) = r(A') = 3$.

Then $\dim(V \cap W) = 1$.

We have $\dim(V+W) + \dim(V \cap W) = \dim(V) + \dim(W)$. Then $\dim(V+W) = \dim(V) + \dim(W) - \dim(V \cap W) = 4$. It happens that $V + W = \mathbb{R}^4$.

(b) Let $B = \begin{bmatrix} 1 & 2 & 7 & 1 & -1 \\ 1 & 1 & 3 & 1 & 0 \end{bmatrix}$, and $C = \begin{bmatrix} 3 & 2 & 5 & -1 & 9 \\ 1 & -1 & -5 & 2 & 0 \end{bmatrix}$.

Define $V = \mathcal{N}(B)$ and $W = \mathcal{N}(C)$.

We have $\dim(V) + r(B) = 5$. Note that $r(B) = r_{row}(B) = 2$. Then $\dim(V) = 3$.

We have $\dim(W) + r(C) = 5$. Note that $r(C) = r_{row}(C) = 2$. Then $\dim(W) = 3$.

Define $A = \left[\begin{array}{c} B \\ \hline C \end{array}\right]$. Note that $\mathcal{N}(A) = \mathcal{N}(B) \cap \mathcal{N}(C) = V \cap W$.

Then $\dim(V \cap W) + r(A) = 5$. We determine the value of $r(A)$:

- We obtain the reduced row-echelon form A' which is row-equivalent to A by applying a sequence of row operations to A:

$$A \longrightarrow \cdots \longrightarrow \begin{bmatrix} 1 & 0 & -1 & 0 & 3 \\ 0 & 1 & 4 & 0 & -1 \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = A'$$

  We see that $r(A) = r(A') = 3$.

Then $\dim(V \cap W) = 2$.

We have $\dim(V+W) + \dim(V \cap W) = \dim(V) + \dim(W)$. Then $\dim(V+W) = \dim(V) + \dim(W) - \dim(V \cap W) = 4$.