L01 Linear space, linear combination and linear transformation

1. Linear space (LS)

   (1) Linear space is a set of elements (vectors)
   $C^{m\times n}$ is the collection of all m by n complex matrices; $C^n = C^{n\times 1}$;
   $R^{m\times n}$ is the collection of all m by n real matrices; $R^m = R^{m\times 1}$.

   (2) Addition is defined in linear sapce
   For $A, B \in C^{m\times n}$, $A+B = (a_{ij})_{m\times n} + (b_{ij})_{m\times n} = (a_{ij}+b_{ij})_{m\times n}$; (elements)
   $A+B = (A_1,...,A_n) + (B_1,...,B_n) = (A_1+B_1,...,A_n+B_n)$; (Columns)
   $A+B = (A_{(1)},...,A_{(m)})' + (B_{(1)},...,B_{(m)})' = (A_{(1)}+B_{(1)},...A_{(m)}+B_{(m)})'$; (rows)
   $$\begin{aligned} A+B &= \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{pmatrix} + \begin{pmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \end{pmatrix} \\ &= \begin{pmatrix} A_{11}+B_{11} & A_{12}+B_{12} & A_{13}+B_{13} \\ A_{21}+B_{21} & A_{22}+B_{22} & A_{23}+B_{23} \end{pmatrix}. \text{Blocks} \end{aligned}$$

   (3) Scalar multiplication
   In scalar multiplication αx, α is an element of a field, usually a complex number or a real number.

   (4) The operations meet a set of familiar requirements
   $x+y = y+x$; $\exists 0$, $0+x = x \forall x$; $\forall x\ \exists(-x)$, $(-x)+x = 0$; $\alpha(x+y) = \alpha x + \alpha y$,....
   Ex1: $C^{m\times n}$, $R^{m\times n}$, C^m, R^n are all LSs.

2. Linear combination (LC)

   (1) Linear combination (LC)
   V is a LS. $x_i \in V$, $i = 1,...,k$. α_i, $i = 1,...,k$, are scalars.
   $\alpha_1 x_1 + \cdots + \alpha_k x_k$ is a vector in V called a linear combination (LC) of $x_1,...,x_k$. $\alpha = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_k \end{pmatrix}$ is the coefficient vector of that LC.

   (2) Subspace
   V is a linear space.
   $$\begin{aligned} S \text{ is a subspace of } V &\overset{def}{\iff} S \subset V \text{ and } S \text{ is a linear space} \\ &\iff S \text{ is closed under two linear operations} \\ &\iff S \text{ is closed under linear combinations.} \end{aligned}$$
   $x_1,...,x_k$ are vectors in a LS V. Let S be the collection of all LCs of $x_1,...,x_k$. Then S is a subspace of V called the Span of $x_1,...,x_n$ denoted by $\text{Span}(x_1,...,x_k)$.

   (3) $\text{Span}(A) = \mathcal{C}(A) = L(A)$
   For $A \in C^{m\times n}$ and $x \in C^n$, $Ax = (A_1,...,A_n)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 A_1 + \cdots + x_n A_n$ is a LC of the columns of A. So $\{Ax : x \in C^n\}$ is the span of the columns of A, also called the column space of A, this is the basic linear space associated with A and hence is denoted by
   $$\text{Span}(A) = \mathcal{C}(A) = L(A) = \{Ax : x \in C^n\} \subset C^m \text{ is a subspace of } C^m.$$
   Ex2: Suppose S_1 and S_2 are two subspaces of S. Then $S_1 \cap S_2$ and $S_1 + S_2$ are two subspaces of S.

Proof. Show the last one. Suppose $x = x_1 + x_2 \in S_1 + S_2$ and $y = y_1 + y_2 \in S_1 + S_2$ where $x_1,\, y_1 \in S_1$ and $x_2,\, y_2 \in S_2$. Then $\alpha x + \beta y = (\alpha x_1 + \beta y_1) + (\alpha x_2 + \beta y_2) \in S_1 + S_2$. So $S_1 + S_2$ is closed under LCs. Hence $S_1 + S_2$ is a subspace of S. $\square$.

3. Linear transformation (LT)

   (1) Linear transformation (LT)
   U and V are two LSs, for $x \in V$ $f(x) \in U$ is a linear transformation (LT) if

   $$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y).$$

   Here $V = \text{Domain}(f)$, $\{f(x) \in U : x \in V\} = \text{Rangle}(f)$, and $\{x \in V : f(x) = 0\} = \text{Kernel}(f)$.

   Ex3: For $A \in C^{m\times n}$, with $x \in C^n$ $f(x) = Ax \in C^m$ is a LT from C^n to C^m since

   $$f(\alpha x + \beta y) = A(\alpha x + \beta y) = \alpha Ax + \beta Ay = \alpha f(x) + \beta f(y).$$

   (2) Range and Kernel of a LT
   For the LT in (1), the Range is a subspace of U and the Kernel is a subspace of V.

   Proof. If $y_1,\, y_2 \in \text{Range}(f)$, then $y_1 = f(x_1)$ and $y_2 = f(x_2)$ for some $x_1,\, x_2 \in V$. So

   $$\alpha y_1 + \beta y_2 = \alpha f(x_1) + \beta f(x_2) = f(\alpha x_1 + \beta x_2) \in \text{Range}(f).$$

   Thus $\text{Range}(f)$ is a subspace of U. If $x_1,\, x_2 \in \text{Kernel}(f)$, then $f(x_1) = 0 = f(x_2)$. So $\qquad f(\alpha x_1 + \beta x_2) = \alpha f(x_1) + \beta f(x_2) = 0$, i.e., $\alpha x_1 + \beta x_2 \in \text{Kernel}(f)$. Thus $\text{Kernel}(f)$ is a subspace of V. $\square$

   (3) $\text{Range}(A) = \mathcal{R}(A)$ and $\text{Kernel}(A) = \mathcal{N}(A)$.
   With $A \in C^{m\times n}$ for the LT $y = Ax$,

   $\text{Range}(A) = \mathcal{R}(A) = \{Ax : x \in C^n\} = \text{Span}(A) = \mathcal{C}(A) = L(A)$ is a subspace of C^m.
   $\text{Kernel}(A) = \mathcal{N}(A) = \{x \in C^n : Ax = 0\}$ is a subspace of C^n.

4. Other operations

   (1) Transpose, conjugate, conjugate-transphse
   Transpose of A: $A' = A^T$. $A' = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}' = \begin{pmatrix} A'_{11} & A'_{21} \\ A'_{12} & A'_{22} \end{pmatrix}$. $A' = A \overset{def}{\Longleftrightarrow} A$ is symmetric
   $\overline{A} = \overline{(a_{ij})}_{m\times n} = (\bar{a}_{ij})_{m\times n}$. $\overline{A} = A \overset{def}{\Longleftrightarrow} A$ is real
   $A^* = A^H = \overline{(A')} = (\overline{A})'$. $A^* = A \overset{def}{\Longleftrightarrow} A$ is Hermitian
   Ex4: A real symmetric matrix is Hermitian.

   (2) Multiplication
   Suppose $B \in C^{n\times k}$. Then $AB = A(B_1, ..., B_k) = (AB_1, ..., AB_k)$ is a set of k ordered LCs of the columns of A with coefficient vectors $B_1, ..., B_k$. When the columns of A and the rows of B are divided the same way, $AB = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = \begin{pmatrix} A_{11}B_1 + A_{12}B_2 \\ A_{21}B_1 + A_{22}B_2 \end{pmatrix}$.

   (3) Relations
   $(AB)' = B'A'$; $\overline{AB} = \overline{A}\,\overline{B}$; $(AB)^* = B^*A^*$. For $A \in C^{m\times n}$, $I_m A = A_n = A$.

L02 Rank of a matrix

1. Ranks and dimensions

   (1) Linear dependence (LD)

$$x_1, ..., x_k \text{ are LD} \overset{def}{\iff} \exists\, \alpha_1 x_1 + \cdots + \alpha_k x_k = 0 \text{ with at least one } \alpha_i \neq 0$$
$$\iff \exists\, x_i \text{ that is a LC of others}$$

   $\Rightarrow$: Suppose $\alpha_1 x_1 + \cdots + \alpha_k x_k = 0$ with $\alpha_i \neq 0$. Solve the equation for x_i.
   $\Leftarrow$: Write x_i as a LC of others. Move x_i to the other side of the equation. $\square$

   (2) Linear independence (LI)

$$x_1, ..., x_k \text{ are LI} \overset{def}{\iff} x_1, ..., x_k \text{ are not LD}$$
$$\iff \text{If } \alpha_1 x_1 + \cdots + \alpha_k x_k = 0, \text{ then } \alpha_1 = \cdots = \alpha_k = 0$$
$$\overset{*}{\iff} \text{If } y \text{ is a LC of } x_1, .., x_k, \text{ then the expression is unique.}$$

   $\overset{*}{\Rightarrow}$: If $y = \sum_i \alpha_i x_i$ and $y = \sum_i \beta_i x_k$, then $\sum_i (\alpha_i - \beta_i) x_i = 0$. So $\alpha_i = \beta_i \;\forall\, i$.
   $\overset{*}{\Leftarrow}$: If $\sum_i \alpha_i x_i = 0$, then $\alpha_i = 0 \,\forall\, i$ since $\sum_i 0\, x_i = 0$. $\square$

   (3) Rank
   Suppose $[x_1, ..., x_r] \subset \mathcal{D} \subset V$ where V is a LS.

$$[x_1, ..., x_r] \text{ is a largest set of LI vectors in } \mathcal{D}$$
$$\overset{def}{\iff} x_1, .., x_r \text{ are LI and } x_1, .., x_r, x \text{ are LD } \forall\, x \in \mathcal{D}$$
$$\iff x_1, ...x_r \text{ are LI and } x \text{ is a LC of } x_1, ..., x_r \text{ for all } x \in \mathcal{D}.$$
$$\iff x \text{ can be expressed as a unique LC of } x_1, ..., x_r \text{ for all } x \in \mathcal{D}.$$

   If both $[x_1, ..., x_r]$ and $[y_1, ..., y_s]$ are largest set of LI vectors in $\mathcal{D}$, then $r = s$. This common value is called the rank of $\mathcal{D}$, $\text{rank}(\mathcal{D}) = r$.

   (4) Basis and dimension
   Suppose $[x_1, ..., x_r] \subset \mathcal{D} \subset V$ where V is a LS.

   If $[x_1, ..., x_r]$ is a largest set of LI vectors in V, then $[x_1, ..., x_r]$ is called a basis of V, and r is called the dimension of V, $\dim(V) = r.V$ is a LS,

   (5) Relation
   Suppose $[x_1, ..., x_r] \subset \mathcal{D} \subset V$ where V is a LS.

   If $[x_1, ..., x_r]$ is a largest set of LI vectors in $\mathcal{D}$, then $[x_1, ..., x_r]$ is a basis for $\text{Span}(\mathcal{D})$. So $\text{rank}(\mathcal{D}) = \dim[\text{Span}(\mathcal{D})]$.

   Proof. $[x_1, .., x_r] \subset \mathcal{D} \subset \text{Span}(\mathcal{D})$; $x_1, ..., x_k$ are LI; For $x \in \text{Span}(\mathcal{D})$, x is a LC of vectors in $\mathcal{D}$, and vectors in $\mathcal{D}$ are LCs of $x_1, ..., x_r$, so x is a LC of $x_1, ..., x_r$. $\square$.

Ex1: If one vectors in $[x_1, ..., x_n]$ is 0, then $x_1, ..., x_n$ are LD.
If $x_1, ..., x_n$ are LD, then $x_1, ..., x_n$, x are LD. If $x_1, ..., x_n$ are LI, then $x_1, .., x_{n-1}$ are LI.

Ex2: $\mathcal{D}_1 \subset \mathcal{D}_2 \Longrightarrow \text{rank}(\mathcal{D}_1) \leq \text{rank}(\mathcal{D}_2)$.

2. Matrix ranks

   (1) rank(A)
   The rank of n columns of $A \in C^{m\times n}$ is called the column rank of A
   The rank of m rows of $A \in C^{m\times n}$ is called the row rank of A
   It can be shown that the column rank and row rank of A are equal. This common value is called the rank of A denoted as rank(A)
   Clearly $0 \le \text{rank}(A) \le m$ and $0 \le \text{rank}(A) \le n$.

   (2) $\text{rank}(A) = \dim[\text{Span}(A)] = \dim[\mathcal{C}(A)] = \dim[\mathcal{R}(A)] = \dim[L(A)]$.

   Ex3: (i) $\text{rank}(A') = \text{rank}(A)$.
   (ii) For $x_i \in C^m$, $i = 1, ..., n$, $\sum_i \alpha_i x_i = 0 \Longleftrightarrow \sum_i \overline{\alpha}_i \, \overline{x}_i = 0$ and $\alpha_i = 0 \Longleftrightarrow \overline{\alpha}_i = 0$.
   So $\text{rank}(\overline{A}) = \text{rank}(A)$.
   (iii) Consequently, $\text{rank}(A^*) = \text{rank}(A)$. So A, A', $\overline{A}$ and A^* share the same rank.

3. Two equations on dimensions
   We present two results without proofs.

   (1) If S_1 and S_2 are two subspaces of S, so are $S_1 + S_2 = \{x + y : x \in S_1 \text{ and } y \in S_2\}$ and $S_1 \cap S_2$. Moreover,

   $$\dim(S_1 + S_2) = \dim(S_1) + \dim(S_2) - \dim(S_1 \cap S_2).$$

   (2) For LT f without proof we present

   $$\dim[\text{domain}(f)] = \dim[\text{Kernel}(f)] + \dim[\text{Range}(f)]$$

   Ex4: $L[(A,\, B)] = L(A) + L(B)$.

   Proof. $\subset$**:** $y \in L[(A,\, B)] \Longrightarrow y = (A,\, B)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = Ax_1 + Bx_2 \in L(A) + L(B)$.

   $\supset$**:** $y \in L(A) + L(B) \Longrightarrow y = Ax_1 + Bx_2 = (A,\, B)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in L[(A,\, B)]$. □

   Ex5:
   $$\begin{aligned}\text{rank}[(A,\, B)] &= \dim\{L[(A,\, B)]\} = \dim[L(A) + L(B)] \\ &= \dim[L(A)] + \dim[L(B)] - \dim[L(A) \cap L(B)] \\ &= \text{rank}(A) + \text{rank}(B) - \dim[L(A) \cap L(B)].\end{aligned}$$
   So $\text{rank}[(A,\, B)] = \text{rank}(A) + \text{rank}(B) - \dim[L(A) \cap L(B)] \le \text{rank}(A) + \text{rank}(B)$

   Ex6: With $A \in C^{m\times n}$, $\dim[\mathcal{N}(A)] = n - \text{rank}(A)$.

   Proof. $f(x) = Ax$ is LT with $\text{domain}(f) = C^n$, $\text{Kernel}(f) = \mathcal{N}(A)$ and $\text{Range}(f) = L(A)$. So

   $$n = \dim(C^n) = \dim[\mathcal{N}(A)] + \dim[L(A)] = \dim[\mathcal{N}(A)] + \text{rank}(A).$$

   Thus $\dim[\mathcal{N}(A)] = n - \text{rank}(A)$.