E2 212: Matrix Theory

Topic: Vector Spaces, Subspaces and Rank



Prof. Sudhan Majhi

TA: Misanolla Vamsi

SPWiCOM LAB
Department of Electrical Communication Engineering
Indian Institute of Science, Bangalore, India.

Vector Spaces

A vector space involves four things: two sets V and $\mathbb{F}$, and two algebraic operations called vector addition and scalar multiplication:

- V is a nonempty set of objects called vectors. Although V can be quite general, we will usually consider V to be a set of n-tuples or a set of matrices.
- $\mathbb{F}$ is a scalar field — for us $\mathbb{F}$ is either the field $\mathbb{R}$ of real numbers or the field $\mathbb{C}$ of complex numbers.
- Vector addition (denoted by $\underline{x}+\underline{y}$) is an operation between elements of V.
- Scalar multiplication (denoted by $\alpha\underline{x}$) is an operation between elements of $\mathbb{F}$ and V.

The formal definition of a vector space stipulates how these four things relate to each other.

Vector Space Definition: The set V is called a *vector space over* $\mathbb{F}$ when the vector addition and scalar multiplication operations satisfy the following properties.

($A1$) $\underline{x}+\underline{y} \in V$ for all $\underline{x}, \underline{y} \in V$. This is called the *closure property* for vector addition.

($A2$) $(\underline{x}+\underline{y})+\underline{z} = \underline{x}+(\underline{y}+\underline{z})$ for every $\underline{x}, \underline{y}, \underline{z} \in V$

$(A3)$ $\underline{x} + \underline{y} = \underline{y} + \underline{x}$ for every $\underline{x}, \underline{y} \in V$

$(A4)$ There is an element $0 \in V$ such that $\underline{x} + 0 = \underline{x}$ for every $\underline{x} \in V$.

$(A5)$ For each $\underline{x} \in V$, there is an element $(-\underline{x}) \in V$ such that $\underline{x} + (-\underline{x}) = 0$.

$(M1)$ $\alpha\underline{x} \in V$ for all $\alpha \in \mathbb{F}$ and $\underline{x} \in V$. This is the *closure property* for scalar multiplication.

$(M2)$ $(\alpha\beta)\underline{x} = \alpha(\beta\underline{x})$ for all $\alpha, \beta \in \mathbb{F}$ and all $\underline{x} \in V$.

$(M3)$ $\alpha(\underline{x} + \underline{y}) = \alpha\underline{x} + \alpha\underline{y}$ for every $\alpha \in \mathbb{F}$ and all $\underline{x}, \underline{y} \in V$.

$(M4)$ $(\alpha + \beta)\underline{x} = \alpha\underline{x} + \beta\underline{x}$ for all $\alpha, \beta \in \mathbb{F}$ and every $\underline{x} \in V$.

$(M5)$ $1\underline{x} = \underline{x}$ for every $\underline{x} \in V$.

Example:

- The set $\mathbb{R}^{m\times n}$ of $m \times n$ real matrices is a vector space over $\mathbb{R}$.
- The set $\mathbb{C}^{m\times n}$ of $m \times n$ complex matrices is a vector space over $\mathbb{C}$.

INDIAN INSTITUTE OF SCIENCE
भारतीय विज्ञान संस्थान

- The real coordinate spaces

$$\mathbb{R}^{1\times n} = \left\{(x_1\ x_2\ \cdots\ x_n),\ x_i \in \mathbb{R}\right\} \quad \text{and} \quad \mathbb{R}^{n\times 1} = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},\ x_i \in \mathbb{R} \right\}$$

  are special cases of the preceding example and these will be the objective of most of our attention.

- With function addition and scalar multiplication defined by

$$(f+g)(x) = f(x) + g(x) \qquad \text{and} \qquad (\alpha f)(x) = \alpha f(x),$$

  the following sets are vector spaces over $\mathbb{R}$:

  - The set of functions mapping the interval $[0,1]$ into $\mathbb{R}$.
  - The set of all real-valued continuous functions defined on $[0,1]$.
  - The set of real-valued functions that are differentiable on $[0,1]$.
  - The set of all polynomials with real coefficients.

INDIAN INSTITUTE OF SCIENCE
भारतीय विज्ञान संस्थान

Consider the vector space $\mathbb{R}^2$, and let

$$L = \{(x, y) \mid y = \alpha x\}$$

be a line through the origin. L is a subset of $\mathbb{R}^2$, but L is a special kind of subset because L also satisfies the properties (A1)–(A5) and (M1)–(M5) that define a vector space. This shows that it is possible for one vector space to properly contain other vector spaces.

Subspaces:

Let S be a nonempty subset of a vector space V over $\mathbb{F}$ (symbolically, $S \subseteq V$). If S is also a vector space over $\mathbb{F}$ using the same addition and scalar multiplication operations, then S is said to be a *subspace* of V. It is not necessary to check all 10 of the defining conditions in order to determine if a subset is also a subspace — only the closure conditions (A1) and (M1) need to be considered. That is, a non-empty subset S of a vector space V is a subspace iff

$(A1)$ $\quad \underline{x}, \underline{y} \in S \implies \underline{x} + \underline{y} \in S$

and

$(M1)$ $\quad \underline{x} \in S \implies \alpha \underline{x} \in S$ for all $\alpha \in \mathbb{F}$

INDIAN INSTITUTE OF SCIENCE
भारतीय विज्ञान संस्थान

Proof: If S is a subset of V, then S automatically inherits all of the vector space properties of V except (A1), (A4), (A5) and (M1). However, (A1) together with (M1) implies (A4) and (A5). To prove this, observe that (M1) implies $(-\underline{x}) = (-1)\underline{x} \in S$ for all $\underline{x} \in S$ so that (A5) holds. Since $\underline{x}$ and $(-\underline{x})$ are now both in S, (A1) insures that $\underline{x} + (-\underline{x}) \in S$, and thus $0 \in S$.

Span(S): For a vectors $S = \{\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_r\}$ from a vector space V, the set of all possible linear combinations of the $\underline{v}_i$'s is denoted by

$$\text{span}(S) = \left\{\alpha_1\underline{v}_1 + \alpha_2\underline{v}_2 + \cdots + \alpha_r\underline{v}_r \;\middle|\; \alpha_i \in \mathbb{F}\right\}.$$

Theorem. span(S) is a subspace of V over $\mathbb{F}$.

Proof: For $\underline{x}, \underline{y} \in \text{span}(S)$, and $\alpha \in \mathbb{F}$ we need to show

$$\begin{aligned} &(A1) && \underline{x} + \underline{y} \in \text{span}(S) \\ &(M1) && \alpha\underline{x} \in \text{span}(S) \end{aligned}$$

If $\underline{x}, \underline{y} \in \text{span}(S)$ we can write

$$\underline{x} = \sum_{i=1}^{r} \xi_i\underline{v}_i, \qquad \underline{y} = \sum_{i=1}^{r} \eta_i\underline{v}_i.$$

(A1) $\therefore \quad \underline{x} + \underline{y} = \sum_{i=1}^{r} (\xi_i + \eta_i)\underline{v}_i \in \text{span}(S)$

(M1) $\alpha\underline{x} = \alpha\sum_{i=1}^{r}\xi_i\underline{v}_i = \sum_{i=1}^{r}(\alpha\xi_i)\underline{v}_i \in \text{span}(S)$

Thus $\text{span}(S)$ is a subspace of V over $\mathbb{F}$.

Spanning sets:

- For a set of vectors $S = \{\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_r\}$, the subspace

$$\text{span}(S) = \{\alpha_1\underline{v}_1 + \alpha_2\underline{v}_2 + \cdots + \alpha_r\underline{v}_r\}$$

  generated by forming all linear combinations of vectors from S is called the *space spanned by* S.

- If V is a vector space such that $V = \text{span}(S)$, we say S is a *spanning set* for V. In other words, S spans V whenever each vector in V is a linear combination of vectors from S.

Examples:

1st class

- $S = \left\{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}2\\2\end{pmatrix}\right\}$ spans the line $y = x$ in $\mathbb{R}^2$.
- The unit vectors $\left\{e_1 = \begin{pmatrix}1\\0\\0\end{pmatrix},\ e_2 = \begin{pmatrix}0\\1\\0\end{pmatrix},\ e_3 = \begin{pmatrix}0\\0\\1\end{pmatrix}\right\}$ span $\mathbb{R}^3$.
- The unit vectors $\{e_1, e_2, \ldots, e_n\}$ in $\mathbb{R}^n$ form a spanning set for $\mathbb{R}^n$.
- The finite set $\{1, x, x^2, \ldots, x^n\}$ spans the space of all polynomials such that $\deg p(x) \leq n$, and the infinite set $\{1, x, x^2, \ldots\}$ spans the space of all polynomials.

Sum of Subspace: If X and Y are subspaces of a vector space V, then the sum of X and Y is defined to be the set of all possible sums of vectors from X with vectors from Y. That is,

$$X + Y = \{\underline{x} + \underline{y} \mid \underline{x} \in X \text{ and } \underline{y} \in Y\}$$

Theorem:

1. The sum $X + Y$ is again a subspace of V.
2. If S_x, S_y span X, Y, then $S_x \cup S_y$ spans $X + Y$.

Proof 1: Let us check the two closure properties (A1) and (M1) for $S = X + Y$.
(A1) If $\underline{u}, \underline{v} \in S$ then $\underline{u} = \underline{x}_1 + \underline{y}_1$ and $\underline{v} = \underline{x}_2 + \underline{y}_2$ where $\underline{x}_1, \underline{x}_2 \in X$ and $\underline{y}_1, \underline{y}_2 \in Y$.

$$\underline{u} + \underline{v} = (\underline{x}_1 + \underline{x}_2) + (\underline{y}_1 + \underline{y}_2) \in S$$

(M1) $\underline{x}_1 \in X$ and $\alpha \in \mathbb{F} \Rightarrow \alpha\underline{x}_1 \in X$; similarly $\underline{y}_1 \in Y \Rightarrow \alpha\underline{y}_1 \in Y$.

$$\alpha\underline{u} = \alpha(\underline{x}_1 + \underline{y}_1) = \alpha\underline{x}_1 + \alpha\underline{y}_1 \in S$$

$\therefore$ The sum $X + Y$ is a subspace of V.

Proof 2: Suppose $S_x = \{\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_r\}$, $S_y = \{\underline{y}_1, \underline{y}_2, \ldots, \underline{y}_t\}$.

$$\underline{z} \in \text{span}(S_x \cup S_y) \implies \underline{z} = \sum_{i=1}^{r} \alpha_i \underline{x}_i + \sum_{i=1}^{t} \beta_i \underline{y}_i = \underline{x} + \underline{y} \quad \text{with } \underline{x} \in X,\ \underline{y} \in Y \implies \underline{z} \in X + Y.$$

INDIAN INSTITUTE OF SCIENCE
भारतीय विज्ञान संस्थान

Subspaces and Linear Functions:

For a linear function f mapping $\mathbb{R}^n$ into $\mathbb{R}^m$, let $\mathcal{R}(f)$ denote the range of f. That is, $\mathcal{R}(f) = \{f(\underline{x}) \mid \underline{x} \in \mathbb{R}^n\} \subseteq \mathbb{R}^m$ is the set of all images as $\underline{x}$ varies freely over $\mathbb{R}^n$.

Theorem. The range of every linear function $f : \mathbb{R}^n \to \mathbb{R}^m$ is a subspace of $\mathbb{R}^m$, and every subspace of $\mathbb{R}^m$ is the range of some linear function. For this reason, subspaces of $\mathbb{R}^m$ are sometimes called *linear spaces*.

Proof: If $f : \mathbb{R}^n \to \mathbb{R}^m$ is a linear function, then the range of f is a subspace of $\mathbb{R}^m$ because the closure properties (A1) and (M1) are satisfied. Establish (A1) by showing that $\underline{y}_1, \underline{y}_2 \in \mathcal{R}(f) \Rightarrow \underline{y}_1 + \underline{y}_2 \in \mathcal{R}(f)$. If $\underline{y}_1, \underline{y}_2 \in \mathcal{R}(f)$, then there must be vectors $\underline{x}_1, \underline{x}_2 \in \mathbb{R}^n$ such that $\underline{y}_1 = f(\underline{x}_1)$ and $\underline{y}_2 = f(\underline{x}_2)$, so it follows from the linearity of f that

$$\underline{y}_1 + \underline{y}_2 = f(\underline{x}_1) + f(\underline{x}_2) = f(\underline{x}_1 + \underline{x}_2) \in \mathcal{R}(f)$$

Similarly, establish (M1) by showing that if $\underline{y} \in \mathcal{R}(f)$, then $\alpha\underline{y} \in \mathcal{R}(f)$ for all scalars α by using the definition of range along with the linearity of f to write

$$\underline{y} \in \mathcal{R}(f) \;\Rightarrow\; \underline{y} = f(\underline{x}) \text{ for some } \underline{x} \in \mathbb{R}^n \;\Rightarrow\; \alpha\underline{y} = \alpha f(\underline{x}) = f(\alpha\underline{x}) \in \mathcal{R}(f)$$

Now prove that every subspace V of $\mathbb{R}^m$ is the range of some linear function $f : \mathbb{R}^n \to \mathbb{R}^m$. Suppose that $\{\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n\}$ is a spanning set for V so that

$$V = \left\{\alpha_1\underline{v}_1 + \cdots + \alpha_n\underline{v}_n \;\middle|\; \alpha_i \in \mathbb{R}\right\}$$

INDIAN INSTITUTE OF SCIENCE
भारतीय विज्ञान संस्थान

Stack the $\underline{v}_i$'s as columns in a matrix $A_{m\times n} = (\underline{v}_1|\underline{v}_2|\cdots|\underline{v}_n)$ and put the α_i's in an $n \times 1$ column $\underline{x} = (\alpha_1, \alpha_2, \ldots, \alpha_n)^T$ to write

$$\alpha_1\underline{v}_1 + \cdots + \alpha_n\underline{v}_n = (\underline{v}_1|\underline{v}_2|\cdots|\underline{v}_n)\begin{pmatrix}\alpha_1\\ \vdots \\ \alpha_n\end{pmatrix} = A\underline{x}$$

The function $f(\underline{x}) = A\underline{x}$ is linear, and we have that

$$\mathcal{R}(f) = \left\{A\underline{x} \;\middle|\; \underline{x} \in \mathbb{R}^{n\times 1}\right\} = \left\{\alpha_1\underline{v}_1 + \cdots + \alpha_n\underline{v}_n \;\middle|\; \alpha_i \in \mathbb{R}\right\} = V.$$

Range Space: The range of a matrix $A \in \mathbb{R}^{m\times n}$ is defined to be the subspace $\mathcal{R}(A)$ of $\mathbb{R}^m$ that is generated by the range of $f(\underline{x}) = A\underline{x}$. That is

$$\mathcal{R}(A) = \left\{A\underline{x} \;\middle|\; \underline{x} \in \mathbb{R}^n\right\} \subseteq \mathbb{R}^m.$$

Similarly the range of A^T is the subspace of $\mathbb{R}^n$ defined by

$$\mathcal{R}(A^T) = \left\{A^T\underline{y} \;\middle|\; \underline{y} \in \mathbb{R}^m\right\} \subseteq \mathbb{R}^n.$$

Because $\mathcal{R}(A)$ is the set of all "images" of vectors $\underline{x} \in \mathbb{R}^n$ under transformation by A, some people call $\mathcal{R}(A)$ the *image space* of A.

Column and Row spaces: For $A \in \mathbb{R}^{m\times n}$ the following statements are true.

- $\mathcal{R}(A)$ = the space spanned by the columns of A (column space)
- $\mathcal{R}(A^T)$ = the space spanned by the rows of A (row space)
- $\underline{b} \in \mathcal{R}(A) \iff \underline{b} = A\underline{x}$ for some $\underline{x}$.
- $\underline{a} \in \mathcal{R}(A^T) \iff \underline{a}^T = \underline{y}^T A$ for some $\underline{y}^T$.

Example: Describe $\mathcal{R}(A)$ and $\mathcal{R}(A^T)$ for $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{pmatrix}$.

Sol^n: $\mathcal{R}(A) = \text{span}\{A_{*1}, A_{*2}, A_{*3}\} = \{\alpha_1 A_{*1} + \alpha_2 A_{*2} + \alpha_3 A_{*3}\}$.
But since $A_{*2} = 2A_{*1}$ and $A_{*3} = 3A_{*1}$

$$\implies \mathcal{R}(A) = \text{span}\{A_{*1}\},$$

is the line in $\mathbb{R}^2$ through the origin and the point $(1, 2)$.
Similarly, $\mathcal{R}(A^T) = \text{span}\{A_{1*}, A_{2*}\} = \{\alpha_1 A_{1*} + \alpha_2 A_{2*} \mid \alpha_1, \alpha_2 \in \mathbb{R}\}$. But $A_{2*} = 2A_{1*} \Rightarrow \mathcal{R}(A^T) = \text{span}(A_{1*})$, and this is the line in $\mathbb{R}^3$ through the origin and the point $(1, 2, 3)$.

Equal Ranges: For two matrices A and B of the same shape

- $\mathcal{R}(A^T) = \mathcal{R}(B^T)$ if and only if $A \overset{\text{row}}{\sim} B$.
- $\mathcal{R}(A) = \mathcal{R}(B)$ if and only if $A \overset{\text{col}}{\sim} B$.

Proof: Assume $A \overset{\text{row}}{\sim} B$ so that there exists a non-singular matrix P such that $PA = B$. To see that $\mathcal{R}(A^T) = \mathcal{R}(B^T)$

$$\begin{aligned} \underline{a} \in \mathcal{R}(A^T) &\iff \underline{a}^T = \underline{y}^T A = \underline{y}^T P^{-1} PA \quad \text{for some } \underline{y}^T \\ &\iff \underline{a}^T = \underline{z}^T B \quad \text{for } \underline{z}^T = \underline{y}^T P^{-1} \\ &\iff \underline{a} \in \mathcal{R}(B^T). \end{aligned}$$

Conversely, if $\mathcal{R}(A^T) = \mathcal{R}(B^T)$, then

$$\text{span}\{A_{1*}, A_{2*}, \ldots, A_{m*}\} = \text{span}\{B_{1*}, B_{2*}, \ldots, B_{m*}\}$$

So each row of B is a combination of the rows of A, and vice versa. On the basis of this fact, it can be argued that it is possible to reduce A to B by using only row operations, and thus $A \overset{\text{row}}{\sim} B$. The next one will be proved by replacing A and B with A^T and B^T. 2nd class

Spanning the Row space and range: Let A be an $m \times n$ matrix, and let U be any row echelon form derived from A. Spanning sets for the row and column spaces are as follows.

- The nonzero rows of U span $\mathcal{R}(A^T)$.
- The basic columns in A span $\mathcal{R}(A)$.

Example: Determine spanning sets of $\mathcal{R}(A)$ and $\mathcal{R}(A^T)$, where $A = \begin{pmatrix} 1 & 2 & 2 & 3 \\ 2 & 4 & 1 & 3 \\ 3 & 6 & 1 & 4 \end{pmatrix}$.

Sol^n:

$$A \simeq U = \begin{bmatrix} 1 & 2 & 2 & 3 \\ 0 & 0 & -3 & -3 \\ 0 & 0 & -5 & -5 \end{bmatrix} \simeq \begin{bmatrix} 1 & 2 & 2 & 3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \simeq \begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Basic columns of A (column operations):

$$A \simeq \begin{bmatrix} 1 & 0 & 2 & 2 \\ 2 & 0 & 1 & 1 \\ 3 & 0 & 1 & 1 \end{bmatrix} \simeq \begin{bmatrix} 1 & 0 & 2 & 0 \\ 2 & 0 & 1 & 0 \\ 3 & 0 & 1 & 0 \end{bmatrix}$$

$$\therefore \quad \mathcal{R}(A) = \text{span}\left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \right\} \qquad \mathcal{R}(A^T) = \text{span}\left\{ \begin{pmatrix} 1 \\ 2 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right\}$$

Example: Null space (or Kernel of space):

- For an $m \times n$ matrix A, the set $\mathcal{N}(A) = \{\underline{x}_{n\times 1} \mid A\underline{x} = 0\} \subseteq \mathbb{R}^n$ is called the *nullspace* of A. In other words, $\mathcal{N}(A)$ is simply the set of all solutions to the homogeneous system $A\underline{x} = 0$.
- The set $\mathcal{N}(A^T) = \{\underline{y}_{m\times 1} \mid A^T\underline{y} = 0\} \subseteq \mathbb{R}^m$ is called the *left-hand nullspace* of A because $\mathcal{N}(A^T)$ is the set of all solutions to the left-hand homogeneous system $\underline{y}^T A = 0^T$.

The basic columns in A span $\mathcal{R}(A)$.

Proof: Basic columns in A are in positions $b_1, b_2, \ldots, b_r$ and non-basic columns are in positions $n_1, n_2, \ldots, n_t$.

Then

$$AQ_1 = \big(B_{m\times r},\ N_{m\times t}\big), \qquad Q_1 \text{ is the permutation matrix.}$$

We can annihilate the nonbasic columns in N using elementary column operations.

$$\implies (B, N)\, Q_2 = (B, 0).$$

$$\implies AQ_1Q_2 = AQ = (B, N)\, Q_2 = (B, 0).$$

$$\implies A \overset{\text{col}}{\sim} (B, 0)$$

$$\implies \text{span}(B) = \mathcal{R}(A)$$

INDIAN INSTITUTE OF SCIENCE
भारतीय विज्ञान संस्थान

Example: Determine a spanning set for $\mathcal{N}(A)$, where $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{pmatrix}$.

Sol^n: $\mathcal{N}(A)$ is merely the general solution of $A\underline{x} = 0$.

The echelon form $E_A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \end{pmatrix}$

$$\implies x_1 = -2x_2 - 3x_3 \quad \text{where } x_2 \text{ and } x_3 \text{ are free}$$

$$\therefore \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -2x_2 - 3x_3 \\ x_2 \\ x_3 \end{pmatrix} = x_2 \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix}$$

In other words, $\mathcal{N}(A)$ is the set of all possible linear combinations of the vectors

$$h_1 = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} \quad \text{and} \quad h_2 = \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix}$$

and therefore $\text{span}\{h_1, h_2\} = \mathcal{N}(A)$. For this example $\mathcal{N}(A)$ is the plane in $\mathbb{R}^3$ that passes through the origin and the two points h_1 and h_2.

Four Fundamental Subspaces:

Spanning the Nullspace: To determine a spanning set for $\mathcal{N}(A)$, where $\text{rank}(A_{m\times n}) = r$, row reduce A to a row echelon form U, and solve $U\underline{x} = 0$ for the basic variables in terms of the free variables to produce the general solution of $A\underline{x} = 0$ in the form

$$\underline{x} = x_{f_1} h_1 + x_{f_2} h_2 + \cdots + x_{f_{n-r}} h_{n-r}.$$

By definition, the set $\mathcal{H} = \{h_1, h_2, \ldots, h_{n-r}\}$ spans $\mathcal{N}(A)$. Moreover, it can be proven that $\mathcal{H}$ is unique in the sense that $\mathcal{H}$ is independent of the row echelon form U.

INDIAN INSTITUTE OF SCIENCE
भारतीय विज्ञान संस्थान

Zero Nullspace: If A is an $m \times n$ matrix, then

- $\mathcal{N}(A) = \{0\}$ if and only if $\text{rank}(A) = n$;
- $\mathcal{N}(A^T) = \{0\}$ if and only if $\text{rank}(A) = m$.

Proof: $\underline{x} = 0$ is the only solution of $A\underline{x} = 0$ iff $\text{rank}(A) = n$. Similarly: $A^T\underline{y} = 0$ will have the trivial solution $\underline{y} = 0$ iff $\text{rank}(A^T) = m$.

Left-hand Nullspace: If $\text{rank}(A_{m\times n}) = r$, and if $PA = U$, where P is nonsingular and U is in row echelon form, then the last $m - r$ rows in P span the left-hand nullspace of A. In other words, if $P = \begin{pmatrix} P_1 \\ P_2 \end{pmatrix}$, where P_2 is $(m-r) \times m$, then

$$\mathcal{N}(A^T) = \mathcal{R}(P_2^T). \qquad \text{(proof: page 14)}$$

Example: Determine a spanning set for $\mathcal{N}(A^T)$, where $A = \begin{pmatrix} 1 & 2 & 2 & 3 \\ 2 & 4 & 1 & 3 \\ 3 & 6 & 1 & 4 \end{pmatrix}$.

Sol^n: To find a nonsingular matrix P such that $PA = U$ is in row echelon form, row reduce the augmented matrix $(A \,|\, I)$ to $(U \,|\, P)$. It must be the case that $PA = U$ because P is the product of the elementary matrices corresponding to the elementary row operations used. Since any row echelon form will suffice, we may use Gauss–Jordan reduction to reduce A to E_A as shown below:

$$\left(\begin{array}{cccc|ccc} 1 & 2 & 2 & 3 & 1 & 0 & 0 \\ 2 & 4 & 1 & 3 & 0 & 1 & 0 \\ 3 & 6 & 1 & 4 & 0 & 0 & 1 \end{array}\right) \xrightarrow{R_2 \to R_2 - 2R_1;\ R_3 \to R_3 - 3R_1} \left(\begin{array}{cccc|ccc} 1 & 2 & 2 & 3 & 1 & 0 & 0 \\ 0 & 0 & -3 & -3 & -2 & 1 & 0 \\ 0 & 0 & -5 & -5 & -3 & 0 & 1 \end{array}\right)$$

$$\xrightarrow{R_2 \to R_2/(-3)} \left(\begin{array}{cccc|ccc} 1 & 2 & 2 & 3 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & \frac{2}{3} & -\frac{1}{3} & 0 \\ 0 & 0 & -5 & -5 & -3 & 0 & 1 \end{array}\right) \xrightarrow{R_3 \to R_3 + 5R_2;\ R_1 \to R_1 - 2R_2} \left(\begin{array}{cccc|ccc} 1 & 2 & 0 & 1 & -\frac{1}{3} & \frac{2}{3} & 0 \\ 0 & 0 & 1 & 1 & \frac{2}{3} & -\frac{1}{3} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{3} & -\frac{5}{3} & 1 \end{array}\right)$$

$$\implies \mathcal{N}(A^T) = \text{span}\left\{ \begin{pmatrix} \frac{1}{3} \\ -\frac{5}{3} \\ 1 \end{pmatrix} \right\} \qquad \text{(last row of } P = P_2\text{)}$$

Equal Nullspaces: For two matrices A and B of the same shape

- $\mathcal{N}(A) = \mathcal{N}(B)$ if and only if $A \overset{\text{row}}{\sim} B$
- $\mathcal{N}(A^T) = \mathcal{N}(B^T)$ if and only if $A \overset{\text{col}}{\sim} B$

Left hand nullspaces proof: If $U = \begin{pmatrix} C \\ 0 \end{pmatrix}$, where $C_{r\times n}$, then $PA = U$ implies $P_2 A = 0$, and this says $\mathcal{R}(P_2^T) \subseteq \mathcal{N}(A^T)$. To show equality, demonstrate containment in the opposite direction by arguing that every vector in $\mathcal{N}(A^T)$ must also be in $\mathcal{R}(P_2^T)$. Suppose $\underline{y}^T \in \mathcal{N}(A^T)$, and let $P^{-1} = (Q_1 \; Q_2)$ to conclude that

$$0 = \underline{y}^T A = \underline{y}^T P^{-1} U = \underline{y}^T Q_1 C \;\Rightarrow\; 0 = \underline{y}^T Q_1$$

because $\mathcal{N}(C^T) = \{0\}$. Now observe that $PP^{-1} = I = P^{-1}P$ insures $P_1 Q_1 = I_r$ and $Q_1 P_1 = I_m - Q_2 P_2$, so

$$\begin{aligned} 0 = \underline{y}^T Q_1 \Rightarrow 0 &= \underline{y}^T Q_1 P_1 = \underline{y}^T (I - Q_2 P_2) \\ \Rightarrow \underline{y}^T &= \underline{y}^T Q_2 P_2 = (\underline{y}^T Q_2) P_2 \\ \Rightarrow \underline{y} &\in \mathcal{R}(P_2^T). \end{aligned}$$

Proof (of Equal Nullspaces): If $\mathcal{N}(A^T) = \mathcal{N}(B^T) \Rightarrow \mathcal{R}(P_2^T) = \mathcal{N}(B^T)$, and hence $P_2 B = 0$. But this means the columns of B are in $\mathcal{N}(P_2)$, that is $\mathcal{R}(B) \subseteq \mathcal{N}(P_2) = \mathcal{R}(A)$ (by using the example on page 15). If A is replaced by B, the result is that $\mathcal{R}(A) \subseteq \mathcal{R}(B)$. Thus we may conclude that $\mathcal{R}(A) = \mathcal{R}(B)$. The second claim may be obtained by replacing A and B by A^T and B^T in above. 3rd class

INDIAN INSTITUTE OF SCIENCE
भारतीय विज्ञान संस्थान

Example: Suppose $\text{rank}(A_{m\times n}) = r$, and let $P = \begin{pmatrix} P_1 \\ P_2 \end{pmatrix}$ be a non-singular matrix such that $PA = U = \begin{pmatrix} C_{r\times n} \\ 0 \end{pmatrix}$, where U is in row echelon form. Prove $\mathcal{R}(A) = \mathcal{N}(P_2)$.

Sol^n: The strategy is to first prove $\mathcal{R}(A) \subseteq \mathcal{N}(P_2)$ and then show the reverse inclusion $\mathcal{N}(P_2) \subseteq \mathcal{R}(A)$. The equation $PA = U$ implies $P_2A = 0$, so all columns of A are in $\mathcal{N}(P_2)$, and thus $\mathcal{R}(A) \subseteq \mathcal{N}(P_2)$. To show inclusion in the opposite direction, suppose $\underline{b} \in \mathcal{N}(P_2)$ so that

$$P\underline{b} = \begin{pmatrix} P_1 \\ P_2 \end{pmatrix} \underline{b} = \begin{pmatrix} P_1\underline{b} \\ P_2\underline{b} \end{pmatrix} = \begin{pmatrix} \underline{d}_{r\times 1} \\ 0 \end{pmatrix}$$

Consequently, $P(A \,|\, \underline{b}) = (PA \,|\, P\underline{b}) = \begin{pmatrix} C & \underline{d} \\ 0 & 0 \end{pmatrix}$, and this implies

$$\text{rank}[A \,|\, \underline{b}] = r = \text{rank}(A).$$

This means the system $A\underline{x} = \underline{b}$ is consistent, and thus $\underline{b} \in \mathcal{R}(A)$. Therefore, $\mathcal{N}(P_2) \subseteq \mathcal{R}(A)$, and we may conclude that $\mathcal{N}(P_2) = \mathcal{R}(A)$.

Summary:

- The range or column space: $\mathcal{R}(A) = \{A\underline{x}\} \subseteq \mathbb{R}^m$
- The row space or left-hand range: $\mathcal{R}(A^T) = \{A^T\underline{y}\} \subseteq \mathbb{R}^n$
- The nullspace: $\mathcal{N}(A) = \{\underline{x} \mid A\underline{x} = 0\} \subseteq \mathbb{R}^n$
- The left-hand nullspace: $\mathcal{N}(A^T) = \{\underline{b} \mid A^T\underline{y} = 0\} \subseteq \mathbb{R}^m$

For $PA = U$ and suppose $\text{rank}(A) = r$.

- Spanning set for $\mathcal{R}(A)$ = the basic columns in A.
- Spanning set for $\mathcal{R}(A^T)$ = the non-zero rows in U.
- Spanning set for $\mathcal{N}(A)$ = the h_i's in the general solution of $A\underline{x} = 0$.
- Spanning set for $\mathcal{N}(A^T)$ = the last $m - r$ rows of P.

Linear Independence: A set of vectors $S = \{\underline{u}_1, \underline{u}_2, \ldots, \underline{u}_n\}$ is said to be a *linearly independent* set whenever the only solution for the scalars α_i in the homogeneous equation

$$\alpha_1 \underline{u}_1 + \alpha_2 \underline{u}_2 + \cdots + \alpha_n \underline{u}_n = 0$$

is the trivial solution $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$.

Basis: A linearly independent spanning set for a vector space V is called a *basis* for V.

Characterizations of a Basis: Let V be a subspace of $\mathbb{R}^m$, and let $\mathcal{B} = \{\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_n\} \subseteq V$. The following statements are equivalent.

- $\mathcal{B}$ is a basis for V.
- $\mathcal{B}$ is a minimal spanning set for V.
- $\mathcal{B}$ is a maximal linearly independent subset of V.

Dimension: The dimension of a vector space V is defined to be

$$\begin{aligned} \dim V &= \text{number of vectors in any basis for } V. \\ &= \text{number of vectors in any minimal spanning set for } V. \\ &= \text{number of vectors in any maximal independent subset of } V. \end{aligned}$$

Subspace Dimension: For vector spaces $\mathcal{M}$ and $\mathcal{N}$ such that $\mathcal{M} \subseteq \mathcal{N}$, the following statements are true.

- $\dim \mathcal{M} \leq \dim \mathcal{N}$.
- If $\dim \mathcal{M} = \dim \mathcal{N}$, then $\mathcal{M} = \mathcal{N}$.

Proof:

(1) Let $\dim \mathcal{M} = m$ and $\dim \mathcal{N} = n$. Now if $m > n$, then there exists a L.I. subset of $\mathcal{N}$ (or basis of $\mathcal{M}$) containing more than n vectors. This is impossible as $\dim(\mathcal{N}) = n$.

(2) If $m = n$ but $\mathcal{M} \neq \mathcal{N}$, then $\exists$ a vector $\underline{x} \in \mathcal{N}$ but $\underline{x} \notin \mathcal{M}$. If $\mathcal{B}$ is a basis for $\mathcal{M}$, then $\underline{x} \notin \text{span}(\mathcal{B})$. The extension set $\mathcal{E} = \mathcal{B} \cup \{\underline{x}\}$ is l.i. and containing $m+1 = n+1$ vectors, which is impossible because $\dim \mathcal{N} = n$.

Hence $m = n \Rightarrow \mathcal{M} = \mathcal{N}$.

Fundamental Subspaces:

For an $m \times n$ matrix of real numbers such that $\text{rank}(A) = r$,

- $\dim \mathcal{R}(A) = r$
- $\dim \mathcal{N}(A) = n - r$
- $\dim \mathcal{R}(A^T) = r$
- $\dim \mathcal{N}(A^T) = m - r$

If $PA = U$ and $\mathcal{H}$ be the set of h_i's appearing in the general solution of $A\underline{x} = 0$

- The basic columns of A form a basis for $\mathcal{R}(A)$.
- The nonzero rows of U form a basis for $\mathcal{R}(A^T)$.
- The set $\mathcal{H}$ is a basis for $\mathcal{N}(A)$.
- The last $m - r$ rows of P form a basis for $\mathcal{N}(A^T)$.

Bases: A linearly independent spanning set for a vector space V is called a basis for V.

Dimension of a Sum: If X and Y are subspaces of a vector space V, then

$$\dim(X+Y) = \dim X + \dim Y - \dim(X \cap Y)$$

Proof: Let say $S = \{\underline{z}_1, \underline{z}_2, \ldots, \underline{z}_t\}$ be a basis for $X \cap Y \Rightarrow S \subseteq X$ and $S \subseteq Y$.
$\exists$ extension vectors $\{\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_m\}$ and $\{\underline{y}_1, \underline{y}_2, \ldots, \underline{y}_n\}$ such that

$$\begin{aligned} \mathcal{B}_x &= \{\underline{z}_1, \underline{z}_2, \ldots, \underline{z}_t, \underline{x}_1, \ldots, \underline{x}_m\} = \text{a basis for } X \\ \text{and} \quad \mathcal{B}_y &= \{\underline{z}_1, \underline{z}_2, \ldots, \underline{z}_t, \underline{y}_1, \underline{y}_2, \ldots, \underline{y}_n\} = \text{a basis for } Y \end{aligned}$$

We know $\mathcal{B} = \mathcal{B}_x \cup \mathcal{B}_y$ spans $X+Y$. We show that $\mathcal{B}$ is linearly independent, i.e.

$$\sum_{i=1}^{t} \alpha_i \underline{z}_i + \sum_{j=1}^{m} \beta_j \underline{x}_j + \sum_{k=1}^{n} \gamma_k \underline{y}_k = 0. \qquad (\circledast)$$

Then

$$\sum_{k=1}^{n} \gamma_k \underline{y}_k = -\Big(\sum_{i=1}^{t} \alpha_i \underline{z}_i + \sum_{j=1}^{m} \beta_j \underline{x}_j\Big) \in X \;\Rightarrow\; \sum_{k=1}^{n} \gamma_k \underline{y}_k \in X \cap Y$$

$\exists\ \delta_i \in \mathbb{F}$ such that

$$\sum_{k=1}^{n} \gamma_k \underline{y}_k = \sum_{i=1}^{t} \delta_i \underline{z}_i \quad \text{or} \quad \sum_{k=1}^{n} \gamma_k \underline{y}_k - \sum_{i=1}^{t} \delta_i \underline{z}_i = 0 \;\Rightarrow\; \gamma_k = 0 = \delta_i \;\; \forall\, i,k \quad \text{as } \mathcal{B}_y \text{ is a basis.}$$

$$\implies \sum_{i=1}^{t} \alpha_i \underline{z}_i + \sum_{j=1}^{m} \beta_j \underline{x}_j = 0$$

But $\mathcal{B}_x$ is also an independent set $\Rightarrow \alpha_i = \beta_j = 0 \ \ \forall\, i, j$.

Thus the only solution of $(\circledast)$ is the trivial solution. Thus $\mathcal{B}$ is linearly independent and a basis of $X + Y$

$$\begin{aligned} \dim(X+Y) = t + m + n &= (t+m) + (t+n) - t \\ &= \dim X + \dim Y - \dim(X \cap Y) \end{aligned}$$

Example: Show that $\text{rank}(A+B) \leq \text{rank}(A) + \text{rank}(B)$.

Sol^n: Observe that $\mathcal{R}(A+B) \subseteq \mathcal{R}(A) + \mathcal{R}(B)$ because if $\underline{b} \in \mathcal{R}(A+B)$, then there exists $\underline{x}$ such that

$$\underline{b} = (A+B)\underline{x} = A\underline{x} + B\underline{x} \in \mathcal{R}(A) + \mathcal{R}(B)$$

$$\begin{aligned} \therefore \quad \text{rank}(A+B) = \dim \mathcal{R}(A+B) &\leq \dim\big(\mathcal{R}(A) + \mathcal{R}(B)\big) \\ &= \dim \mathcal{R}(A) + \dim \mathcal{R}(B) - \dim\big(\mathcal{R}(A) \cap \mathcal{R}(B)\big) \\ &\leq \dim \mathcal{R}(A) + \dim \mathcal{R}(B) \\ &= \text{rank}(A) + \text{rank}(B). \end{aligned}$$

Rank of a product: If A is $m \times n$ and B is $n \times p$, then

$$\text{rank}(AB) = \text{rank}(B) - \dim\big(\mathcal{N}(A) \cap \mathcal{R}(B)\big). \textit{4thclass}$$

Proof: Let $S = \{\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_s\}$ be a basis for $\mathcal{N}(A) \cap \mathcal{R}(B)$. If $\dim \mathcal{R}(B) = s + t$

Then $\mathcal{B} = \{\underline{x}_1, \ldots, \underline{x}_s, \underline{z}_1, \ldots, \underline{z}_t\}$ is a basis for $\mathcal{R}(B)$ for $\exists$ extension set $S_{ext} = \{\underline{z}_1, \underline{z}_2, \ldots, \underline{z}_t\}$.
We claim $\mathcal{T} = \{A\underline{z}_1, A\underline{z}_2, \ldots, A\underline{z}_t\}$ is a basis for $\mathcal{R}(AB)$.
$\mathcal{T}$ spans $\mathcal{R}(AB)$ because if $\underline{b} \in \mathcal{R}(AB)$, then $\underline{b} = AB\underline{y}$ for some $\underline{y}$, but $B\underline{y} \in \mathcal{R}(B)$ implies $B\underline{y} = \sum_{i=1}^{s} \xi_i \underline{x}_i + \sum_{i=1}^{t} \eta_i \underline{z}_i$. So

$$\underline{b} = A\Big(\sum_{i=1}^{s} \xi_i \underline{x}_i + \sum_{i=1}^{t} \eta_i \underline{z}_i\Big) = \sum_{i=1}^{s} \xi_i A\underline{x}_i + \sum_{i=1}^{t} \eta_i A\underline{z}_i$$
$$= \sum_{i=1}^{t} \eta_i A\underline{z}_i \qquad (A\underline{x}_i = 0 \text{ as } S \subseteq \mathcal{N}(A))$$

$\mathcal{T}$ is l.i. $\Rightarrow$

$$0 = \sum_{i=1}^{t} \alpha_i A\underline{z}_i = A\sum_{i=1}^{t} \alpha_i \underline{z}_i \;\Rightarrow\; \sum_{i=1}^{t} \alpha_i \underline{z}_i \in \mathcal{N}(A) \cap \mathcal{R}(B)$$
$$\Rightarrow \sum_{i=1}^{t} \alpha_i \underline{z}_i = \sum_{j=1}^{s} \beta_j \underline{x}_j \quad \text{or} \quad \sum_{i=1}^{t} \alpha_i \underline{z}_i - \sum_{j=1}^{s} \beta_j \underline{x}_j = 0$$

A trivial solution $\alpha_i = \beta_j = 0$ for all i, j as $\mathcal{B}$ is an independent set.
$\Rightarrow$ Thus $\mathcal{T}$ is a basis for $\mathcal{R}(AB)$, so $t = \dim \mathcal{R}(AB) = \text{rank}(AB)$ and hence

$$\text{rank}(B) = \dim \mathcal{R}(B) = s + t = \dim\big(\mathcal{N}(A) \cap \mathcal{R}(B)\big) + \text{rank}(AB).$$

INDIAN INSTITUTE OF SCIENCE
भारतीय विज्ञान संस्थान

Bounds on the Rank of a product: If A is $m \times n$ and B is $n \times p$ then

① $\text{rank}(AB) \leq \min\{\text{rank}(A),\ \text{rank}(B)\}$

① $\text{rank}(A) + \text{rank}(B) - n \leq \text{rank}(AB)$

and

$$\text{rank}(A^TA) = \text{rank}(A) = \text{rank}(AA^T) = \text{rank}(A^T).$$

Proof:

②

$$\begin{aligned}\mathcal{N}(A) \cap \mathcal{R}(B) \subseteq \mathcal{N}(A) &\Rightarrow \dim\big(\mathcal{N}(A) \cap \mathcal{R}(B)\big) \leq \dim \mathcal{N}(A) = n - \text{rank}(A) \\ &\Rightarrow n - \text{rank}(A) \geq \dim\big(\mathcal{N}(A) \cap \mathcal{R}(B)\big)\end{aligned}$$

Lower bound:

$$\begin{aligned}\text{rank}(AB) &= \text{rank}(B) - \dim\big(\mathcal{N}(A) \cap \mathcal{R}(B)\big) \\ &\geq \text{rank}(B) + \text{rank}(A) - n.\end{aligned}$$

①

$$\begin{aligned}&\text{rank}(AB) = \text{rank}(B) - \dim\big(\mathcal{N}(A) \cap \mathcal{R}(B)\big) \leq \text{rank}(B) \\ &\text{rank}(AB) = \text{rank}(AB)^T = \text{rank}(B^TA^T) \leq \text{rank}(A^T) = \text{rank}(A) \\ &\implies \text{rank}(AB) \leq \min\big(\text{rank}(A),\ \text{rank}(B)\big)\end{aligned}$$