Linear Algebra

Diego ROJAS

June 4, 2022

1 Introduction

Linear algebra is a discipline of mathematics that deals with linear equations, vectors, matrices and their transformations.

2 Systems of linear equations

Definition: **(Linear system):** A collection of one or more equations involving the same variables.

Rule: For any linear system of the following form:

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned} \tag{1}$$

There exists a shorthand notation:

$$\begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} x_1 + \begin{bmatrix} a_{12} \\ \vdots \\ a_{m2} \end{bmatrix} x_1 + \cdots + \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix} x_n = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} \tag{2}$$

Which can be further compacted into the form $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ like so:

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} \tag{3}$$

Or using the *augmented matrix* notation $[\boldsymbol{A}|\boldsymbol{b}]$

$$\left[\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right] \tag{4}$$

3 Matrices

Definition: **(Matrix):** With $m, n \in \mathbb{N}$ a real-valued (m, n) matrix $\boldsymbol{A}$ is an $m \cdot n$ tuple of elements a_{ij}. $\mathbb{R}^{m \times n}$ is the set of all real valued matrices.

Rule: For matrices $\boldsymbol{A} \in \mathbb{R}^{m \times n}, \boldsymbol{B} \in \mathbb{R}^{n \times k}$, the elements c_{ij} of the product $\boldsymbol{C} = \boldsymbol{AB} \in \mathbb{R}^{m \times k}$ are computed as:

$$c_{ij} = \sum_{l=1}^{n} = a_{il} b_{lj}, \text{ where } i = 1, \ldots, m, \text{ and } j = 1, \ldots, k. \tag{5}$$

Rule: For matrices $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{m \times n}$, addition is defined as element-wise.

Rule: Element wise multiplication of $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{m \times n}$ is called a *Hadamard product*.

Definition: **(Identity matrix):** We define an identity matrix in $\mathbb{R}^{n \times n}$ as:

$$I_n := \begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} \in \mathbb{R}^{n \times n} \tag{6}$$

Definition: **(Inverse matrix):** The inverse of a matrix $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ denoted $\boldsymbol{A}^{-1}$ is a unique matrix $\boldsymbol{B} \in \mathbb{R}^{n\times n}$ given $\boldsymbol{AB} = \boldsymbol{I}_n = \boldsymbol{BA}$. If there exists an inverse for a matrix $\boldsymbol{A}$ then it is called *regular/invertible/nonsingular*, otherwise *singular/non-invertible*.

Definition: **(Transpose):** For $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ the matrix $\boldsymbol{B} \in \mathbb{R}^{n\times m}$ with $b_{ij} = a_{ji}$ is called the transpose of $\boldsymbol{A}$. We write $\boldsymbol{B} = \boldsymbol{A}^\top$.

4 Solving systems of linear equations

Rule: When solving a system of linear equations of the form $\boldsymbol{Ax} = \boldsymbol{b}$ we're looking to find scalars $x_1, \ldots, x_n$ such that $\sum_{i=1}^n x_i \boldsymbol{c}_i = \boldsymbol{b}$. A *particular solution* of a linear system is a solution to the equation with particular, arbitrary values. A *general solution* of a system of linear equations can be obtained using the following steps

- Use gaussian elimination to find a particular solution of $\boldsymbol{Ax} = \boldsymbol{b}$.
- Find all solutions of $\boldsymbol{Ax} = \boldsymbol{0}$.
- Combine these solutions onto a set.

Definition: **(Elementary transformations):** Consists of three rule that can be applied on a linear system of the form $\boldsymbol{Ax} = \boldsymbol{b}$.

- Exchange of two equations (swapping rows in $[\boldsymbol{A}|\boldsymbol{b}]$)
- Multiplication of an equation with a constant $\lambda \in \mathbb{R}\backslash\{0\}$
- Addition of two equations (rows of $[\boldsymbol{A}|\boldsymbol{b}]$)

We use $\rightsquigarrow$ to indicate an elementary transformation.

Definition: **(Pivot | Leading coefficient):** The *leading coefficient* of a row (first nonzero number from the left) is called the *pivot* and is always strictly to the right of the pivot of the row above it.

Definition: **(Row echelon form):** A matrix being in row echelon form means that Gaussian elimination has operated on the rows or columns of that matrix such that:

- All rows consisting of only zeroes are at the bottom.
- The leading coefficient of a non-zero row is always strictly to the right of the leading coefficient of the row above it.

Definition: **(Reduced row echelon form):** An equation is in *reduced row echelon form* (or *row cannonical form*) if every pivot is 1, the pivot is the only nonzero entry in its column (each pivot's column is a unit vector).

Rule: The variables corresponding to the pivots in the row-echelon form are called *basic variables* and the other variables are *free variables*.

Rule: To obtain the inverse $\boldsymbol{A}^{-1}$ of a matrix $\boldsymbol{A}$ we can use Gaussian elimination to bring it into reduced row echelon form such as the desired inverse is given as its right hand side.

$$[\boldsymbol{A}|\boldsymbol{I}_n] \rightsquigarrow \cdots \rightsquigarrow [\boldsymbol{I}_n|\boldsymbol{A}^{-1}] \tag{7}$$

5 Vector spaces

Definition: **(Group):** Given a set $\mathcal{G}$ and an operation $\otimes$ defined on $\mathcal{G}$. $G := (\mathcal{G}, \otimes)$ is a group if the following holds:

- Closure of $\mathcal{G}$ under $\otimes$: $\forall x, y \in \mathcal{G} : x \otimes y \in \mathcal{G}$
- Associativity: $\forall x, y, z \in \mathcal{G} : (x \otimes y) \otimes z = x \otimes (y \otimes z)$
- Neutral element: $\exists e \in \mathcal{G}\ \forall x \in \mathcal{G} : x \otimes e = x$
- Inverse element: $\forall x \in \mathcal{G}\ \exists y \in \mathcal{G} : x \otimes y = e$ where e is the neutral element.

Definition: **(Vector space):** A real-valued vector space $V = (\mathcal{V}, +, \cdot)$ is a set $\mathcal{V}$ with two operations $+$ and $\cdot$ defined where the following holds:

- $(\mathcal{V}, +)$ is an abelian group.
- Distributivity:
  - $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathcal{V}\ \forall \lambda \in \mathbb{R} : \lambda(\boldsymbol{x} + \boldsymbol{y}) = \lambda \boldsymbol{x} + \lambda \boldsymbol{y}$
  - $\forall \boldsymbol{x} \in \mathcal{V}\ \forall \lambda, \psi \in \mathbb{R} : (\lambda + \psi)\boldsymbol{x} = \lambda \boldsymbol{x} + \psi \boldsymbol{y}$
- Associativity (outer operation)
- Neutral element (outer operation)

Definition: **(Vector subspace):** Let $V = (\mathcal{V}, +, \cdot)$ be a vector space and $\mathcal{U} \subseteq \mathcal{V}, \mathcal{U} \neq \emptyset$. Then $U = (\mathcal{U}, +, \cdot)$ is a *vector subspace* (or *linear subspace*) if U is a vector space restricted to $\mathcal{U} \times \mathcal{U}$ and $\mathcal{U} \times \mathbb{R}$. We write $U \subseteq V$.

Rule: Given $U \subseteq V$, the following properties of V are passed to U.

- Abelian group, distributivity, associativity and neutral element properties.
- $\mathcal{U} \neq \emptyset$, in particular: $\mathbf{0} \in \mathcal{U}$.
- Closure of U:
  - $\forall \lambda \in \mathbb{R}\ \forall \boldsymbol{x} \in \mathcal{U} : \lambda \boldsymbol{x} \in \mathcal{U}$.
  - $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathcal{U} : \boldsymbol{x} + \boldsymbol{y} \in \mathcal{U}$.

Rule: Every subspace $U \subseteq (\mathbb{R}^n, +, \cdot)$ is the solution space of a homogeneous system of linear equations $\boldsymbol{A}\boldsymbol{x} = \mathbf{0}$ for $\boldsymbol{x} \in \mathbb{R}^n$

6 Linear independence

Definition: **(Linear combination):** Consider a vector space V and a finite number of vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k \in V$. Then every $\boldsymbol{v} \in V$ of the form:

$$\boldsymbol{v} = \sum_{i=1}^{k} \lambda_i \boldsymbol{x}_i \in V \tag{8}$$

with $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ is a linear combination of the vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$.

Definition: **(Linear independence):** Given a vector space V and $k \in \mathbb{N}$. Vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k \in V$ are said to be linearly independent if there exists no non-trivial solution to $\mathbf{0} = \sum_{i=1}^{k} \lambda_i \boldsymbol{x}_i$ with at least one $\lambda_i \neq 0$. Otherwise they're linearly dependent.

Rule:

- If at least one of the vector $\boldsymbol{x}_i$ is $\mathbf{0}$ then they are linearly dependent.
- The vectors $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k : \boldsymbol{x}_i \neq \mathbf{0}, i = 1, \ldots, k\}, k \geqslant 2$ are linearly dependent, if and only if, at least one of them is a linear combination of the others.
- To check whether $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k \in V$ are linearly independent we can use Gaussian elimation: write all vectors as columns of a matrix $\boldsymbol{A}$ and perform Gaussian elimination until the matrix is in row-echelon form.
  - Non-pivot columns can be expressed as linear combinations of vectors on their left.
  - Pivot columns are linearly independent from vectors on their left.

  If all columns are pivots, the vectors are linearly independent.

Rule: Given m linear combinations over k linearly independent vectors $\boldsymbol{b}_1, \dots, \boldsymbol{b}_k \in V$.

$$\begin{aligned} \boldsymbol{x}_1 &= \sum_{i=1}^{k} \lambda_{i1} \boldsymbol{b}_i \\ &\vdots \\ \boldsymbol{x}_m &= \sum_{i=1}^{k} \lambda_{im} \boldsymbol{b}_i \end{aligned} \tag{9}$$

We can write, with $\boldsymbol{B} = [\boldsymbol{b}_1, \dots, \boldsymbol{b}_k]$, the following:

$$\boldsymbol{x}_j = \boldsymbol{B}\boldsymbol{\lambda}_j, \quad \boldsymbol{\lambda}_j = \begin{bmatrix} \lambda_1 j \\ \vdots \\ \lambda_k j \end{bmatrix}, \quad j = 1, \dots, m \tag{10}$$

We can test whether $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m$ are linearly independent using:

$$\sum_{j=1}^{m} \psi_j \boldsymbol{x}_j = \sum_{j=1}^{m} \psi_j \boldsymbol{B}\boldsymbol{\lambda}_j = \boldsymbol{B} \sum_{j=1}^{m} \psi_j \boldsymbol{\lambda}_j \tag{11}$$

Which means that $\{\boldsymbol{x}_1, \dots, \boldsymbol{x}_k\}$ is linearly independent if the column vectors $\{\boldsymbol{\lambda}_1, \dots, \boldsymbol{\lambda}_m\}$ are linearly independent.

Rule: In a vector space V, m linear combinations of $\boldsymbol{b}_1, \dots, \boldsymbol{b}_k$ are linearly independent if $m > k$.

7 Basis and rank

In a vector space V, we are interested in a set of vectors $\mathcal{A}$ that posess the property that any vector $\boldsymbol{v} \in V$ can be obtained through a linear combination of vectors in $\mathcal{A}$.

7.1 Generating Set and Basis

Definition: **(Generating Set and Span):** Given a set of vectors $\mathcal{A} = \{\boldsymbol{x}_1, \dots, \boldsymbol{x}_k\} \subseteq \mathcal{V}$. If every vector $\boldsymbol{v} \in \mathcal{V}$ can be expressed as a linear combination of vectors in $\mathcal{A}$, $\mathcal{A}$ is called a *generating set* of V. The set of all linear combinations of vectors in $\mathcal{A}$ is called the span of $\mathcal{A}$. We write $V = \text{span}[\mathcal{A}]$ or $V = \text{span}[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k]$.

Definition: **(Basis):** Consider a vector space $V = (\mathcal{V}, +, \cdot)$ and $\mathcal{A} \subseteq \mathcal{V}$. A generating set $\mathcal{A}$ of $\mathcal{V}$ is called *minimal* if there exists no smaller set $\tilde{\mathcal{A}} \subsetneq \mathcal{A} \subseteq \mathcal{V}$ that spans $\mathcal{V}$. Every linearly independent generating set of V is minimal and is called a *basis* of V.

Rule: Given $\mathcal{B} \subseteq \mathcal{V}, \mathcal{B} \neq \emptyset$. Then the following statements are equivalent:

- $\mathcal{B}$ is a basis of V
- $\mathcal{B}$ is a minimal generating set.
- $\mathcal{B}$ is a maximal linearly independent set of vectors in V such that adding any vector to the set $\mathcal{B}$ would make it linearly dependent.
- Every vector $\boldsymbol{x} \in V$ is a linear combination of vectors from $\mathcal{B}$, and every combination is unique, i.e., with

$$\boldsymbol{x} = \sum_{i=1}^{k} \lambda_i \boldsymbol{b}_i \tag{12}$$

Rule: Every vector space V posseses a basis B. There can be many bases for V, all have the same number of elements, the *basis vectors*.

Rule: A basis of a subspace $U = \text{span}[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k] \subseteq \mathbb{R}^n$ can be found by:

- Writing the spanning vectors as columns of a matrix $\boldsymbol{A}$
- Determining the row-echelon form of $\boldsymbol{A}$
- The spanning vectors associated with the pivots columns are a basis of U.

7.2 Rank

Definition: **(Rank):** The number of linearly independent columns of a matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ equals the number of linearly independent rows and is called the *rank*. We write $\text{rk}(\boldsymbol{A})$.

Rule:

- $\text{rk}(\boldsymbol{A}) = \text{rk}(\boldsymbol{A}^\top)$ .i.e column rank equals row rank.
- The columns of $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ span a subspace $U \subseteq \mathbb{R}^m$ with $\dim(U) = \text{rk}(\boldsymbol{A})$. This subspace is called the *image* or *range*. A basis of U can be found by applying Gaussian elimination to $\boldsymbol{A}$. Same goes for the rows, given a subspace $W \in \mathbb{R}^n$.
- For all $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ it holds that $\boldsymbol{A}$ is regular (invertible) if and only $\text{rk}(\boldsymbol{A}) = n$.
- For all $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ and all $\boldsymbol{b} \in \mathbb{R}^m$ it holds that the linear equation system $\boldsymbol{Ax} = \boldsymbol{b}$ can be solved only and if $\text{rk}(\boldsymbol{A}) = \text{rk}(\boldsymbol{A}|\boldsymbol{b})$
- A matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ is said to have *full-rank* if $\text{rk}(\boldsymbol{A}) = \min(m, n)$. Otherwise, it is *rank-deficient*.

8 Linear mappings

Definition: **(Linear mapping):** For vector spaces V, W, a mapping $\Phi : V \to W$ is called a *linear mapping* (or *linear transformation* or *vector space homomorphism*). Linear transformations are generally represented as matrices.

$$\begin{aligned} \Phi(\boldsymbol{x} + \boldsymbol{y}) &= \Phi(\boldsymbol{x}) + \Phi(\boldsymbol{y}) \\ \Phi(\lambda \boldsymbol{x}) &= \lambda \Phi(\boldsymbol{x}) \end{aligned} \tag{13}$$

Definition: **(Injective, Surjective, Bijective):** Consider a mapping $\Phi : \mathcal{V} \to \mathcal{W}$ where $\mathcal{V}, \mathcal{W}$ can be arbitrary sets. Then Φ is called

- *Injective* if $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathcal{V} : \Phi(\boldsymbol{x}) = \Phi(\boldsymbol{y}) \implies \boldsymbol{x} = \boldsymbol{y}$.

- *Surjective* if $\Phi(\mathcal{V}) = \mathcal{W}$.
- *Bijective* if it is injective and surjective. Which means there exists a mapping Ψ so that $\Psi \circ \Phi(\boldsymbol{x}) = \boldsymbol{x}$. This mapping Ψ is called the inverse of Φ and denoted Φ^{-1}.

Rule:

- *Isomorphism* $\Phi : V \to W$ linear and bijective.
- *Endomorphism* $\Phi : V \to V$ linear.
- *Automorphism* $\Phi : V \to V$ linear and bijective.
- We define $\mathrm{id}_V : V \to V, \boldsymbol{x} \mapsto \boldsymbol{x}$ as the *identity mapping* or *identity automorphism* in V.

Rule: *Finite dimensional vector spaces* $\mathcal{V}$ and $\mathcal{W}$ are *isomorphic* ($\mathcal{V} \to \mathcal{W}$ linear and bijective) if and only if $\dim(\mathcal{V}) = \dim(\mathcal{W})$.

Definition: **(Ordered basis):** Matrix representation of the basis vectors for an n dimensional vector space $\mathcal{V}$. We write:

$$\begin{aligned} \text{Ordered basis notation: } & B = (\boldsymbol{b}_1, \dots, \boldsymbol{b}_n) \\ \text{Matrix notation: } & \boldsymbol{B} = [\boldsymbol{b}_1, \dots, \boldsymbol{b}_n] \\ \text{Unordered basis notation: } & \mathcal{B} = \{\boldsymbol{b}_1, \dots, \boldsymbol{b}_n\} \end{aligned} \tag{14}$$

Definition: **(Coordinates):** Given an ordered basis $B = (\boldsymbol{b}_1, \dots, \boldsymbol{b}_n)$ The *coordinate vector/coordinate representation* of $\boldsymbol{x} \in V$ is given as:

$$\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \in \mathbb{R}^n \tag{15}$$

Where

$$\boldsymbol{x} = \alpha_1 \boldsymbol{b}_1 + \dots + \alpha_n \boldsymbol{b}_n \tag{16}$$

Definition: **(Transformation Matrix):** Considering vector spaces V, W with corresponding (ordered) bases $B = (\boldsymbol{b}_1, \dots, \boldsymbol{b}_n)$ and $C = (\boldsymbol{c}_1, \dots, \boldsymbol{c}_m)$ and a linear mapping $\Phi : V \to W$. For $j \in \{1, \dots, n\}$:

$$\Phi(\boldsymbol{b}_j) = \alpha_{1j}\boldsymbol{c}_1 + \cdots + \alpha_{mj}\boldsymbol{c}_m \tag{17}$$

is the unique representation of $\Phi(\boldsymbol{b}_j)$ with respect to C. Then we call the $m \times n$ matrix $\boldsymbol{A}_\Phi$, whose elements are given by

$$A_\Phi(i, j) = \alpha_{ij} \tag{18}$$

the transformation matrix of Φ (with respect to B, V, C, W). Therefore, if $\hat{x}$ is the coordinate vector of $x \in \mathcal{V}$ with respect to B and $\hat{y}$ the coordinate vector of $y = \Phi(\boldsymbol{x}) \in \mathcal{W}$ with respect to C, then

$$\hat{\boldsymbol{y}} = \boldsymbol{A}_\Phi \hat{\boldsymbol{x}} \tag{19}$$

Definition: **(Basis change):** For a linear mapping $\Phi : V \to W$, the ordered bases $B, \tilde{B}$ of V of n elements and $C, \tilde{C}$ of W of m elements and the transformation matrix $\boldsymbol{A}_\Phi$ of Φ with respect to B and C, $\tilde{\boldsymbol{A}}_\Phi$ with respect to $\tilde{B}$ and $\tilde{C}$ is given as

$$\tilde{\boldsymbol{A}}_\Phi = \boldsymbol{T}^{-1}\boldsymbol{A}_\Phi \boldsymbol{S} \tag{20}$$

where $\boldsymbol{S} \in \mathbb{R}^{n \times n}$ is the transformation matrix of id_V and $\boldsymbol{T} \in \mathbb{R}^{m \times m}$ is the transformation matrix of id_W.

Definition: **(Equivalence):** Two matrices $\boldsymbol{A}, \tilde{\boldsymbol{A}} \in \mathbb{R}^{m \times n}$ are *equivalent* if there exists regular matrices $\boldsymbol{S} \in \mathbb{R}^{n \times n}$ and $\boldsymbol{T} \in \mathbb{R}^{m \times m}$ such that $\tilde{\boldsymbol{A}} = \boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{S}$.

Definition: **(Similarity):** Two matrices $\boldsymbol{A}, \tilde{\boldsymbol{A}} \in \mathbb{R}^{n \times n}$ are *similar* if there exists a regular matrix $\boldsymbol{S} \in \mathbb{R}^{n \times n}$ such that $\tilde{\boldsymbol{A}} = \boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S}$.

Rule: Considering vector spaces V, W, X, the linear mapping $\Phi : V \to W$ and the linear mapping $\Psi : W \to X$. We know that the linear mapping $\Psi \circ \Phi : V \to X$ is linear and we can write $\boldsymbol{A}_{\Psi \circ \Phi} = \boldsymbol{A}_\Psi \boldsymbol{A}_\Phi$. Therefore

- $\boldsymbol{A}_\Phi$ is the transformation matrix of a linear mapping $\Phi_{CB} : V \to W$

- $\tilde{\boldsymbol{A}}_\Phi$ is the transformation matrix of a linear mapping $\Phi_{\tilde{C}\tilde{B}} : V \to W$
- $\boldsymbol{S}$ is the transformation matrix of a linear mapping $\Psi_{B\tilde{B}} : V \to V$
- $\boldsymbol{T}$ is the transformation matrix of a linear mapping $\Xi_{C\tilde{C}} : W \to W$

Definition: **(Dimension):** The *dimension* of a vector space V is the number of basis vectors of V.

Definition: **(Kernel):** The kernel $\ker(\Phi)$ is the set of vectors that Φ maps to the neutral element $\mathbf{0}_W \in W$. It is defined as

$$\ker(\Phi) := \Phi^{-1}(\mathbf{0}_W) = \{\boldsymbol{v} \in V : \Phi(\boldsymbol{v}) = \mathbf{0}_W\} \tag{21}$$

Definition: **(Image):** The *image* or *range* of a linear transformation $\Phi : V \to W$ is the set of vectors in W that can be "reached" by Φ from any vector in V. It is defined as

$$\mathrm{Im}(\Phi) := \Phi(V) = \{\boldsymbol{w} \in W | \exists \boldsymbol{v} \in V : \Phi(\boldsymbol{v}) = \boldsymbol{w}\} \tag{22}$$

Rule:

- The null space is never empty because $\Phi(\mathbf{0}_V) = \mathbf{0}_W$ always holds.
- $\mathrm{Im}(\Phi) \subseteq W$ is a subspace of W and $\ker(\Phi) \in V$ is a subspace of V.
- Φ is injective if and only if $\ker(\Phi) = \mathbf{0}$ (one to one mapping between V and W with respect to Φ).

Definition: **(Column space):** The image is equivalent to the span of the columns of the transformation matrix. This is called the *column space*.

Definition: **(Null space):** The *null space* $\ker(\Phi)$ is the general solution to the homogeneous linear system $\boldsymbol{A}\boldsymbol{x} = \mathbf{0}$.

Rule: The *rank-nullity theorem* also reffered to as the *fundamental theorem of linear mappings* states that

$$\dim(\ker(\Phi)) + \dim(\mathrm{Im}(\Phi)) = \dim(V) \tag{23}$$

Therefore the following is true

- If $\dim(\mathrm{Im}(\Phi)) < \dim(V)$, then $\ker(\Phi)$ is non-trivial, i.e., the kernel contains more than $\mathbf{0}_V$ and $\dim(\ker(\Phi)) \geqslant 1$.
- If $\dim(\mathrm{Im}(\Phi)) < \dim(V)$, then the system of linear equations $\boldsymbol{A}_\Phi \boldsymbol{x} = \mathbf{0}$ has infinitely many solutions.
- If $\dim(V) = \dim(W)$ then the Φ is bijective since $\mathrm{Im}(\Phi) \subseteq W$.

9 Affine Subspaces

Definition: **(Affine subspace):** Let V be a vector space and $U \subseteq V$ a subset. Then

$$\begin{aligned} L = \boldsymbol{x}_0 + U &:= \{\boldsymbol{x}_0 + \boldsymbol{u} : \boldsymbol{u} \in U\} \\ &= \{\boldsymbol{v} \in V | \exists \boldsymbol{u} \in U : \boldsymbol{v} + \boldsymbol{x}_0 = \boldsymbol{u}\} \subseteq V \end{aligned} \tag{24}$$

is called *affine subspace* or *linear manifold* of V. U is called *direction* or *direction space*, and $\boldsymbol{x}_0$ is called *support point*.

Definition: **(Affine mapping):** For two vector spaces V, W, a linear mapping $\Phi : V \to W$ and $\boldsymbol{a} \in W$, the mapping

$$\begin{aligned} \phi :\ & V \to W \\ & \boldsymbol{x} \mapsto \boldsymbol{a} + \Phi(\boldsymbol{x}) \end{aligned} \tag{25}$$

is an affine mapping from V to W. $\boldsymbol{a}$ is called the *translation vector* of ϕ.