Lesson 7: Vector spaces, subspaces, left null space

- New concepts:
  - Vector space
  - Linear mappings
  - Vector subspace
  - Left null space

Up to now we have taken vectors as belonging to some set $\boldsymbol{v} \in \mathcal{V}$, such as $\boldsymbol{b} \in \mathbb{R}^m$, without any special restrictions on $\mathcal{V}$. In applications we require $\mathcal{V}$ to have additional structure, such as ensuring that the sum of two vectors stays within the set.

Definition. $(\mathcal{V}, \mathcal{S}, +)$ *is a vector space if for any* $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathcal{V}$, *and any* $\alpha, \beta \in \mathcal{S}$, *with* $\mathcal{S}$ *a scalar field, the following properties hold:*

Closed. $\boldsymbol{u} + \boldsymbol{v} \in \mathcal{V}$

Associativity. $\boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w}) = (\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w}$

Null element. $\exists \boldsymbol{0} \in \mathcal{V}$ *such that* $\boldsymbol{u} + \boldsymbol{0} = \boldsymbol{u}$

Inverse element. $\exists (-\boldsymbol{u})$ *such that* $\boldsymbol{u} + (-\boldsymbol{u}) = \boldsymbol{0}$

Commutativity. $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$

Distributivity over scalar addition. $(\alpha + \beta)\boldsymbol{u} = \alpha\boldsymbol{u} + \beta\boldsymbol{v}$

Distributivity over vector addition. $\alpha(\boldsymbol{u} + \boldsymbol{v}) = \alpha\boldsymbol{u} + \alpha\boldsymbol{v}$

Scalar identity. $1 \in \mathcal{S} \Rightarrow 1\boldsymbol{u} = \boldsymbol{u}$

Recall the definition of a function $f: X \mapsto Y$: a relation between two sets X, Y such that to any element in $x \in X$ there is associated a single element in Y denoted as $f(x) \in Y$.

Definition. *A linear mapping is a function between two vector spaces $\mathcal{X}, \mathcal{Y}$ over the same scalar field $\mathcal{S}$, $f: \mathcal{X} \mapsto \mathcal{Y}$, with the property that $\forall \boldsymbol{u}, \boldsymbol{v} \in \mathcal{X}$, $\forall \alpha, \beta \in \mathcal{S}$*

$$f(\alpha \boldsymbol{u} + \beta \boldsymbol{v}) = \alpha f(\boldsymbol{u}) + \beta f(\boldsymbol{v})$$

A matrix $A \in \mathbb{R}^{m \times n}$ is a linear mapping from the vector space $(\mathbb{R}^n, \mathbb{R}, +)$ to the vector space $(\mathbb{R}^m, \mathbb{R}, +)$

$$\boldsymbol{A}: \mathbb{R}^n \mapsto \mathbb{R}^m$$

$$\boldsymbol{x} \in R^n, \boldsymbol{b} \in \mathbb{R}^m, \boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}$$

Note: matrices are linear mappings between vector spaces for which we now the number of components necessary to specify vectors in the domain, codomain.

The most commonly used vector spaces are $\mathbb{R}^m$ for some $m \in \mathbb{N}$. For example $\mathbb{R}^3$ is used as a model for three-dimensional Euclidean space in mechanics. Within three-dimensional space we recognize entities such as lines, planes that require fewer than 3 numbers to specify a position on the line or within the plane. This concept is fomalized by the definition

Definition. *$\mathcal{U}$ is a vector subspace of vector space $\mathcal{V}$ over the same scalar field $\mathcal{S}$ if for any $\boldsymbol{u}, \boldsymbol{v} \in \mathcal{U}$, any $\alpha, \beta \in \mathcal{S}$*

Inclusion. $\boldsymbol{u} \in \mathcal{V}$ *(a vector in the subspace is also in the enclosing vector space)*

Closed. $\alpha \boldsymbol{u} + \beta \boldsymbol{v} \in \mathcal{U}$ *(linear combinations of subspace vectors stay within the subspace)*

We denote the vector subspace relation by $\mathcal{U} \leq \mathcal{V}$ as in $\mathbb{R} \leq \mathbb{R}^3$ stating that the real line is a subspace of three-dimensional space.

Column, null spaces as subspaces of codomain, domain of $\boldsymbol{A}$

Recall definitions of column space, null space of $\boldsymbol{A} \in \mathbb{R}^{m \times n}$

$$
\begin{aligned}
C(\boldsymbol{A}) &= \{\boldsymbol{b} \in \mathbb{R}^m \mid \exists \boldsymbol{x} \in \mathbb{R}^n \text{ such that } \boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}\} \subseteq \mathbb{R}^m \\
N(\boldsymbol{A}) &= \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}\} \subseteq \mathbb{R}^n
\end{aligned}
$$

Note that $C(\boldsymbol{A}) \subseteq \mathbb{R}^m$, $N(\boldsymbol{A}) \subseteq \mathbb{R}^n$ means that $C(\boldsymbol{A})$, $N(\boldsymbol{A})$ are subsets of $\mathbb{R}^m$, $\mathbb{R}^n$ respectively. In fact, we can make a stronger statement, that they are vector subspaces

$$
C(\boldsymbol{A}) \leq \mathbb{R}^m, N(\boldsymbol{A}) \leq \mathbb{R}^n
$$

Proof. Let $\boldsymbol{u}, \boldsymbol{v} \in C(\boldsymbol{A})$, $\alpha, \beta \in \mathcal{S}$. By defintion of $C(\boldsymbol{A})$ there exist $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ such that $\boldsymbol{u} = \boldsymbol{A}\boldsymbol{x}$ and $\boldsymbol{v} = \boldsymbol{A}\boldsymbol{y}$. Using vector space properties

$$
\alpha \boldsymbol{u} + \beta \boldsymbol{v} = \alpha \boldsymbol{A}\boldsymbol{x} + \beta \boldsymbol{A}\boldsymbol{y} = \boldsymbol{A}(\alpha \boldsymbol{x} + \beta \boldsymbol{y}),
$$

hence $\alpha \boldsymbol{u} + \beta \boldsymbol{v} \in C(\boldsymbol{A})$ (it is obtained as the image through the linear mapping $\boldsymbol{A}$ of $\alpha \boldsymbol{x} + \beta \boldsymbol{y}$)

Orthogonality implies linear independence

- Recall that if $\boldsymbol{u}^T\boldsymbol{v} = 0$, with $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^m$ then $\boldsymbol{u} \perp \boldsymbol{v}$ (orthogonal)

Proposition. *If $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n \in \mathbb{R}^m$ are non-zero ($\boldsymbol{u}_i \neq \boldsymbol{0}$) and pairwise orthogonal, $\boldsymbol{u}_i^T\boldsymbol{u}_j = 0$ for $i \neq j$ then they form a linearly independent set of vectors.*

Proof. *Consider the equation equating the linear combination $c_1\boldsymbol{u}_1 + \ldots + c_n\boldsymbol{u}_n$ to the zero vector*

$$c_1\boldsymbol{u}_1 + \ldots + c_n\boldsymbol{u}_n = \boldsymbol{0} \tag{1}$$

Multiply on the left by $\boldsymbol{u}_i^T$ and use orthogonality to obtain $c_i = 0$ for $i = 1, \ldots, n$. The only solution to (1) is $c_1 = c_2 = \ldots = c_n = 0$, hence $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\}$ is a linearly independent set.

Left null space

- We've seen that $C(\boldsymbol{A}) \leq \mathbb{R}^m$, the row space of a matrix is a subspace of $\mathbb{R}^m$
- Typically it is not the entire space $\mathbb{R}^m$
- What's the missing part?
- These would be vectors within $\mathbb{R}^m$ not reachable by linear combinations of the columns of $\boldsymbol{A}$.
- One way to characterize this is to consider vectors $\boldsymbol{y} \in \mathbb{R}^m$ orthogonal to the columns of $\boldsymbol{A}$

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{a}_1 & \dots & \boldsymbol{a}_n \end{pmatrix}, \boldsymbol{a}_j^T \boldsymbol{y} = 0$$

- The above can more compactly be written as $\boldsymbol{A}^T \boldsymbol{y} = \boldsymbol{0}$ with

$$\boldsymbol{A}^T = \begin{pmatrix} \boldsymbol{a}_1^T \\ \vdots \\ \boldsymbol{a}_n^T \end{pmatrix}$$

- Define the *left null space* of $\boldsymbol{A}$ as $N(\boldsymbol{A}^T) \subseteq \mathbb{R}^m$