- Recall: to each matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ associate four fundamental subspaces:
  1. *Column space*, $C(\boldsymbol{A}) = \{\boldsymbol{b} \in \mathbb{R}^m | \, \exists \boldsymbol{x} \in \mathbb{R}^n \, \text{such that} \, \boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}\} \subseteq \mathbb{R}^m$, the part of $\mathbb{R}^m$ *reachable* by linear combination of columns of $\boldsymbol{A}$
  2. *Left null space*, $N(\boldsymbol{A}^T) = \{\boldsymbol{y} \in \mathbb{R}^m | \, \boldsymbol{A}^T \boldsymbol{y} = 0\} \subseteq \mathbb{R}^m$, the part of $\mathbb{R}^m$ *not reachable* by linear combination of columns of $\boldsymbol{A}$
  3. *Row space*, $R(\boldsymbol{A}) = C(\boldsymbol{A}^T) = \{\boldsymbol{c} \in \mathbb{R}^n | \, \exists \boldsymbol{y} \in \mathbb{R}^m \, \text{such that} \, \boldsymbol{c} = \boldsymbol{A}^T \boldsymbol{y}\} \subseteq \mathbb{R}^n$, the part of $\mathbb{R}^m$ *reachable* by linear combination of rows of $\boldsymbol{A}$
  4. *Null space*, $N(\boldsymbol{A}) = \{\boldsymbol{x} \in \mathbb{R}^n | \, \boldsymbol{A}\boldsymbol{x} = 0\} \subseteq \mathbb{R}^n$, the part of $\mathbb{R}^n$ *not reachable* by linear combination of rows of $\boldsymbol{A}$

The fundamental theorem of linear algebra (FTLA) states that for $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, encoding a linear mapping $\boldsymbol{b} = \boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}$, $\boldsymbol{f}: \mathbb{R}^n \to \mathbb{R}^m$

$$C(\boldsymbol{A}), N(\boldsymbol{A}^T) \leq \mathbb{R}^m, C(\boldsymbol{A}) \perp N(\boldsymbol{A}^T), \; C(\boldsymbol{A}) \cap N(\boldsymbol{A}^T) = \{\mathbf{0}\}, \; C(\boldsymbol{A}) \oplus N(\boldsymbol{A}^T) = \mathbb{R}^m$$
$$C(\boldsymbol{A}^T), N(\boldsymbol{A}) \leq \mathbb{R}^n, C(\boldsymbol{A}^T) \perp N(\boldsymbol{A}), \; C(\boldsymbol{A}^T) \cap N(\boldsymbol{A}) = \{\mathbf{0}\}, \; C(\boldsymbol{A}^T) \oplus N(\boldsymbol{A}) = \mathbb{R}^n$$

- The theorem provides an orthogonal decomposition of the domain and codomain of $\boldsymbol{f}$
- It specifies the condition on $\boldsymbol{x}$ for $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, namely $\boldsymbol{b} \in C(\boldsymbol{A})$

Lemma 1. *Let $\mathcal{U}, \mathcal{V}$, be subspaces of vector space $\mathcal{W}$. Then $\mathcal{W} = \mathcal{U} \oplus \mathcal{V}$ if and only if*

i. $\mathcal{W} = \mathcal{U} + \mathcal{V}$, *and*

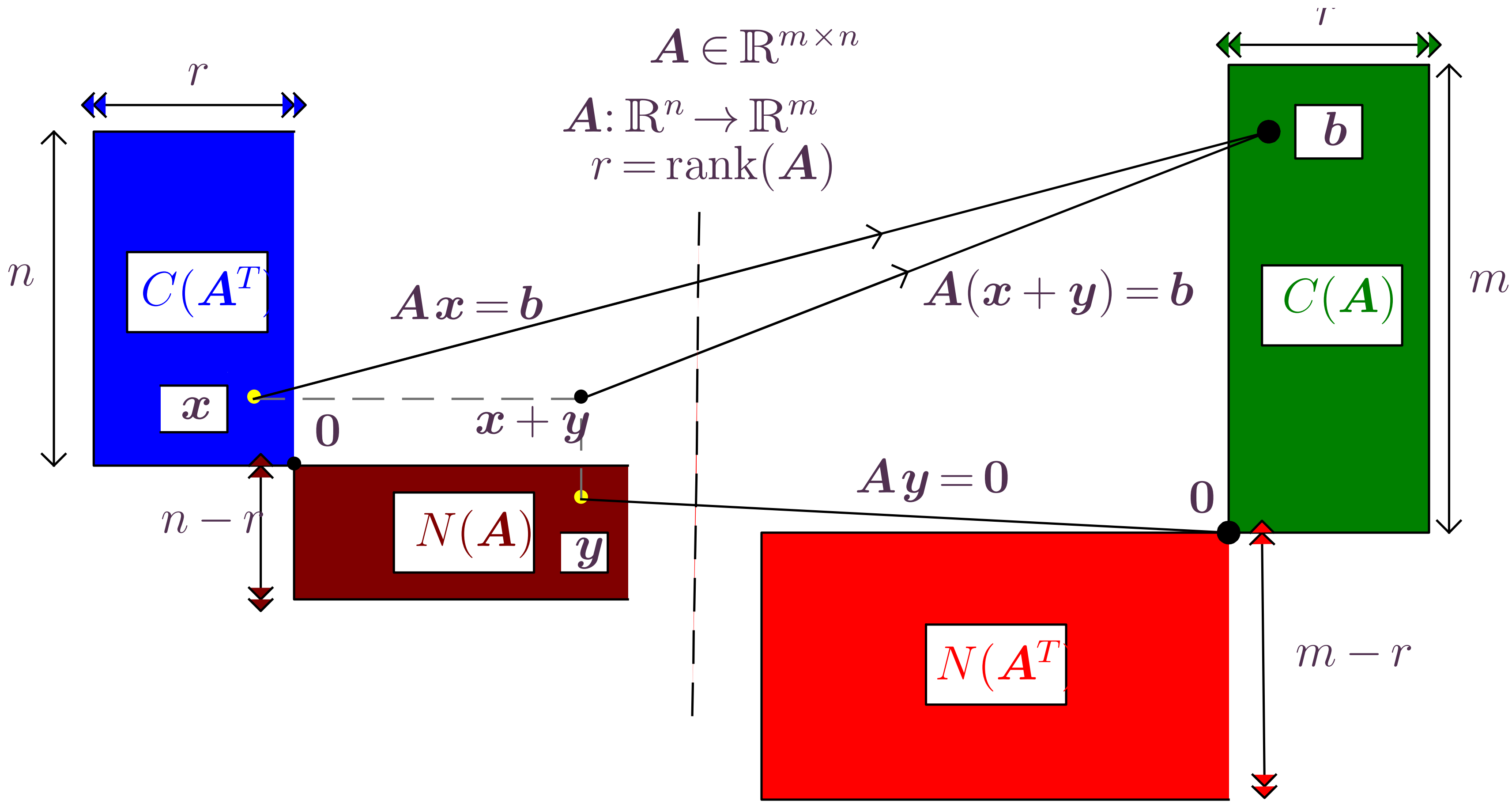
ii. $\mathcal{U} \cap \mathcal{V} = \{\mathbf{0}\}$.

Lemma 2. *Orthogonal complements of $\mathbb{R}^m$ ($m \in \mathbb{N}$, finite), $\mathcal{U}, \mathcal{V} \leq \mathbb{R}^m$, $\mathcal{U} = \mathcal{V}^\perp$, $\mathcal{V} = \mathcal{U}^\perp$, form a direct sum $\mathcal{U} \oplus \mathcal{V} = \mathbb{R}^m$. (proved after discussion of Gram-Schmidt procedure)*

Proof. $\mathcal{W}=\mathcal{U}\oplus\mathcal{V}\Rightarrow\mathcal{W}=\mathcal{U}+\mathcal{V}$ by definition of direct sum, sum of vector subspaces. To prove that $\mathcal{W}=\mathcal{U}\oplus\mathcal{V}\Rightarrow\mathcal{U}\cap\mathcal{V}=\{\mathbf{0}\}$, consider $\boldsymbol{w}\in\mathcal{U}\cap\mathcal{V}$. Since $\boldsymbol{w}\in\mathcal{U}$ and $\boldsymbol{w}\in\mathcal{V}$ write

$$\boldsymbol{w}=\boldsymbol{w}+\mathbf{0}\ \ (\boldsymbol{w}\in\mathcal{U},\mathbf{0}\in\mathcal{V}),\ \ \boldsymbol{w}=0+\boldsymbol{w}\ \ (\mathbf{0}\in\mathcal{U},\boldsymbol{w}\in\mathcal{V}),$$

and since expression $\boldsymbol{w}=\boldsymbol{u}+\boldsymbol{v}$ is unique, it results that $\boldsymbol{w}=\mathbf{0}$. Now assume (i),(ii) and establish an unique decomposition. Assume there might be two decompositions of $\boldsymbol{w}\in\mathcal{W}$, $\boldsymbol{w}=\boldsymbol{u}_1+\boldsymbol{v}_1$, $\boldsymbol{w}=\boldsymbol{u}_2+\boldsymbol{v}_2$, with $\boldsymbol{u}_1,\boldsymbol{u}_2\in\mathcal{U}$, $\boldsymbol{v}_1,\boldsymbol{v}_2\in\mathcal{V}$. Obtain $\boldsymbol{u}_1+\boldsymbol{v}_1=\boldsymbol{u}_2+\boldsymbol{v}_2$, or $\boldsymbol{x}=\boldsymbol{u}_1-\boldsymbol{u}_2=\boldsymbol{v}_2-\boldsymbol{v}_1$. Since $\boldsymbol{x}\in\mathcal{U}$ and $\boldsymbol{x}\in\mathcal{V}$ it results that $\boldsymbol{x}=\mathbf{0}$, and $\boldsymbol{u}_1=\boldsymbol{u}_2$, $\boldsymbol{v}_1=\boldsymbol{v}_2$, i.e., the decomposition is unique. $\square$



$\mathbb{R}^n = C(\boldsymbol{A}^T) \oplus N(\boldsymbol{A})$ usually: $m \geqslant n$ $\mathbb{R}^m = N(\boldsymbol{A}^T) \oplus C(\boldsymbol{A})$

$C(\boldsymbol{A}^T) \perp N(\boldsymbol{A})$ $N(\boldsymbol{A}^T) \perp C(\boldsymbol{A})$

i. $C(\boldsymbol{A}) \leq \mathbb{R}^m$ (column space is vector subspace of codomain of $\boldsymbol{A}: \mathbb{R}^n \to \mathbb{R}^m$)

Proof. Consider arbitrary $\boldsymbol{u}, \boldsymbol{v} \in C(\boldsymbol{A})$, $\alpha, \beta \in \mathbb{R}$. Verify vector subspace properties (Lesson 7 p.4):

i. Inclusion (elements of $C(\boldsymbol{A})$ are in $\mathbb{R}^m$) $\boldsymbol{u} \in \mathbb{R}^m$, yes by definition of $C(\boldsymbol{A}) = \{\boldsymbol{b} \in \mathbb{R}^m | \ \exists \boldsymbol{x} \in \mathbb{R}^n \text{ such that } \boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}\}$. (This immediately results from definitions and will not be shown explicitly in following proofs).

ii. Closed ($\alpha\boldsymbol{u} + \beta\boldsymbol{v} \in C(\boldsymbol{A})$). By definition of $C(\boldsymbol{A})$, $\boldsymbol{u}, \boldsymbol{v} \in C(\boldsymbol{A})$ implies existence of $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ such that $\boldsymbol{u} = \boldsymbol{A}\boldsymbol{x}$, $\boldsymbol{v} = \boldsymbol{A}\boldsymbol{y}$. Compute $\alpha\boldsymbol{u} + \beta\boldsymbol{v} = \alpha(\boldsymbol{A}\boldsymbol{x}) + \beta(\boldsymbol{A}\boldsymbol{y}) = \boldsymbol{A}(\alpha\boldsymbol{x} + \beta\boldsymbol{y})$, and note that since $\alpha\boldsymbol{x} + \beta\boldsymbol{y} \in \mathbb{R}^n$, $\alpha\boldsymbol{u} + \beta\boldsymbol{v} \in C(\boldsymbol{A})$. $\square$

ii. $N(\boldsymbol{A}^T) \leq \mathbb{R}^m$ (left null space is vector subspace of domain of $\boldsymbol{A}: \mathbb{R}^n \to \mathbb{R}^m$)

Proof. Consider arbitrary $\alpha, \beta \in \mathbb{R}$, $\boldsymbol{x}, \boldsymbol{y} \in N(\boldsymbol{A}^T) \Rightarrow \boldsymbol{A}^T\boldsymbol{x} = \boldsymbol{0}$, $\boldsymbol{A}^T\boldsymbol{y} = \boldsymbol{0}$. Compute $\boldsymbol{A}^T(\alpha\boldsymbol{x} + \beta\boldsymbol{y}) = \alpha(\boldsymbol{A}^T\boldsymbol{x}) + \beta(\boldsymbol{A}^T\boldsymbol{y}) = \alpha \cdot \boldsymbol{0} + \beta \cdot \boldsymbol{0} = \boldsymbol{0}$, hence $\alpha\boldsymbol{x} + \beta\boldsymbol{y} \in N(\boldsymbol{A}^T)$ $\square$

iii. $C(\boldsymbol{A}) \perp N(\boldsymbol{A}^T)$ (column space is orthogonal to left null space).

Proof. Consider arbitrary $\boldsymbol{u} \in C(\boldsymbol{A}), \boldsymbol{v} \in N(\boldsymbol{A}^T)$. By definition of $C(\boldsymbol{A})$, $\exists \boldsymbol{x} \in \mathbb{R}^n$ such that $\boldsymbol{u} = \boldsymbol{A}\boldsymbol{x}$, and by definition of $N(\boldsymbol{A}^T)$, $\boldsymbol{A}^T\boldsymbol{v} = \boldsymbol{0}$. Compute $\boldsymbol{u}^T\boldsymbol{v} = (\boldsymbol{A}\boldsymbol{x})^T\boldsymbol{v} = \boldsymbol{x}^T\boldsymbol{A}^T\boldsymbol{v} = \boldsymbol{x}^T(\boldsymbol{A}^T\boldsymbol{v}) = \boldsymbol{x}^T\boldsymbol{0} = \boldsymbol{0}$, hence $\boldsymbol{u} \perp \boldsymbol{v}$ for arbitrary $\boldsymbol{u}, \boldsymbol{v}$, and $C(\boldsymbol{A}) \perp N(\boldsymbol{A}^T)$. □

iv. $C(\boldsymbol{A}) \cap N(\boldsymbol{A}^T) = \{\boldsymbol{0}\}$ ($\boldsymbol{0}$ is the only vector both in $C(\boldsymbol{A})$ and $N(\boldsymbol{A}^T)$).

Proof. (By contradiction, *reductio ad absurdum*). Assume there might be $\boldsymbol{b} \in C(\boldsymbol{A})$ and $b \in N(\boldsymbol{A}^T)$ and $\boldsymbol{b} \neq \boldsymbol{0}$. Since $\boldsymbol{b} \in C(\boldsymbol{A})$, $\exists \boldsymbol{x} \in \mathbb{R}^n$ such that $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}$. Since $\boldsymbol{b} \in N(\boldsymbol{A}^T)$, $\boldsymbol{A}^T\boldsymbol{b} = \boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{x}) = \boldsymbol{0}$. Note that $\boldsymbol{x} \neq 0$ since $\boldsymbol{x} = 0 \Rightarrow \boldsymbol{b} = 0$, contradicting assumptions. Multiply equality $\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ on left by $\boldsymbol{x}^T$,

$$\boldsymbol{x}^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0} \Rightarrow (\boldsymbol{A}\boldsymbol{x})^T(\boldsymbol{A}\boldsymbol{x}) = \boldsymbol{b}^T\boldsymbol{b} = \|\boldsymbol{b}\|^2 = \boldsymbol{0},$$

thereby obtaining $\boldsymbol{b} = \boldsymbol{0}$, using norm property 3 (Lesson 4, p5). Contradiction.

□

v. $C(\boldsymbol{A}) \oplus N(\boldsymbol{A}^T) = \mathbb{R}^m$

Proof. (iii) and (iv) have established that $C(\boldsymbol{A}), N(\boldsymbol{A}^T)$ are orthogonal complements

$$C(\boldsymbol{A}) = N(\boldsymbol{A}^T)^\perp, N(\boldsymbol{A}^T) = C(\boldsymbol{A})^\perp.$$

By Lemma 2 it results that $C(\boldsymbol{A}) \oplus N(\boldsymbol{A}^T) = \mathbb{R}^m$. (Reminder: Proof of Lemma 2 is postponed until discussion of the Gram-Schmidt procedure). □

The remainder of the FTLA is established by considering $\boldsymbol{B} = \boldsymbol{A}^T$, e.g., since it has been established in (v) that $C(\boldsymbol{B}) \oplus N(\boldsymbol{A}^T) = \mathbb{R}^n$, replacing $\boldsymbol{B} = \boldsymbol{A}^T$ yields $C(\boldsymbol{A}^T) \oplus N(\boldsymbol{A}) = \mathbb{R}^m$, etc.

Remark. The great widespread aplicability of linear algebra results in large part due to the complete characterization of the possible solutions to $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ provided by the FTLA and the orthogonal decomposition of the domain and codomain.