FUNDAMENTAL THEOREM OF LINEAR ALGEBRA

SYNOPSIS. Vectors have been introduced as a mathematical object to represent complicated objects. A framework for working with vectors based upon observations of velocity vectors has been introduced. A procedure for obtaining new vectors from some predetermined set has been formally defined as a linear combination, and the matrix-vector product has been introduced to concisely state this operation. Consideration of multiple linear combinations leads to the definition of matrix-matrix products. Functions that transform input vectors into output vectors have been defined, with those that preserve linear combinations playing a distinguished role, the linear mappings. Matrices also arise in the description of linear mappings. Now that the mathematical framework has been established, a natural question to ask is whether it is complete. Could all questions of interest about linear combinations be answered within this framework. The Fundamental Theorem of Linear Algebra (FTLA) gives an affirmative, but non-constructive answer to this question.

1. Partition of linear mapping domain and codomain

A partition of a set S is a collection of subsets $P=\{S_i | S_i \subset P, S_i \neq \emptyset\}$ such that any given element $x \in S$ belongs to only one set in the partition. This is modified when applied to subspaces of a vector space, and a partition of a set of vectors is understood as a collection of subsets such that any vector except $\mathbf{0}$ belongs to only one member of the partition.

Linear mappings between vector spaces $\boldsymbol{f}: U \longrightarrow V$ can be represented by matrices $\boldsymbol{A}$ with columns that are images of the columns of a basis $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots\}$ of U

$$\boldsymbol{A} = [\ \boldsymbol{f}(\boldsymbol{u}_1)\ \ \boldsymbol{f}(\boldsymbol{u}_2)\ \ \ldots\ ].$$

Consider the case of real finite-dimensional domain and co-domain, $\boldsymbol{f}: \mathbb{R}^n \longrightarrow \mathbb{R}^m$, in which case $\boldsymbol{A} \in \mathbb{R}^{m \times n}$,

$$\boldsymbol{A} = [\ \boldsymbol{f}(\boldsymbol{e}_1)\ \ \boldsymbol{f}(\boldsymbol{e}_2)\ \ \ldots\ \ \boldsymbol{f}(\boldsymbol{e}_n)\ ] = [\ \boldsymbol{a}_1\ \ \boldsymbol{a}_2\ \ \ldots\ \ \boldsymbol{a}_n\ ].$$

The column space of $\boldsymbol{A}$ is a vector subspace of the codomain, $C(\boldsymbol{A}) \leq \mathbb{R}^m$, but according to the definition of dimension if $n<m$ there remain non-zero vectors within the codomain that are outside the range of $\boldsymbol{A}$,

$$n<m \Longrightarrow \exists \boldsymbol{v} \in \mathbb{R}^m, \boldsymbol{v} \neq \mathbf{0}, \boldsymbol{v} \notin C(\boldsymbol{A}).$$

All of the non-zero vectors in $N(\boldsymbol{A}^T)$, namely the set of vectors orthogonal to all columns in $\boldsymbol{A}$ fall into this category. The above considerations can be stated as

$$C(\boldsymbol{A}) \leq \mathbb{R}^m,\ \ N(\boldsymbol{A}^T) \leq \mathbb{R}^m,\ \ C(\boldsymbol{A}) \perp N(\boldsymbol{A}^T)\ \ C(\boldsymbol{A}) + N(\boldsymbol{A}^T) \leq \mathbb{R}^m.$$

The question that arises is whether there remain any non-zero vectors in the codomain that are not part of $C(\boldsymbol{A})$ or $N(\boldsymbol{A}^T)$. The fundamental theorem of linear algebra states that there no such vectors, that $C(\boldsymbol{A})$ is the orthogonal complement of $N(\boldsymbol{A}^T)$, and their direct sum covers the entire codomain $C(\boldsymbol{A}) \oplus N(\boldsymbol{A}^T) = \mathbb{R}^m$.

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

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

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

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

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

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 $\boldsymbol{w} \in \mathscr{W}$, $\boldsymbol{w} = \boldsymbol{u}_1 + \boldsymbol{v}_1$, $\boldsymbol{w} = \boldsymbol{u}_2 + \boldsymbol{v}_2$, with $\boldsymbol{u}_1, \boldsymbol{u}_2 \in \mathscr{U}$, $\boldsymbol{v}_1, \boldsymbol{v}_2 \in \mathscr{V}$. 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 $\boldsymbol{x} \in \mathscr{U}$ and $\boldsymbol{x} \in \mathscr{V}$ 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. □

In the vector space $U + V$ the subspaces U, V are said to be orthogonal complements is $U \perp V$, and $U \cap V = \{\mathbf{0}\}$. When $U \leq \mathbb{R}^m$, the orthogonal complement of U is denoted as $U^\perp$, $U \oplus U^\perp = \mathbb{R}^m$.

2. The FTLA

THEOREM. *Given the linear mapping associated with matrix $A \in \mathbb{R}^{m \times n}$ we have:*

1. $C(\boldsymbol{A}) \oplus N(A^T) = \mathbb{R}^m$, *the direct sum of the column space and left null space is the codomain of the mapping*

2. $C(A^T) \oplus N(\boldsymbol{A}) = \mathbb{R}^n$, *the direct sum of the row space and null space is the domain of the mapping*

3. $C(\boldsymbol{A}) \perp N(A^T)$ *and* $C(A) \cap N(A^T) = \{\mathbf{0}\}$, *the column space is orthogonal to the left null space, and they are orthogonal complements of one another,*

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

4. $C(\boldsymbol{A}^T) \perp N(\boldsymbol{A})$ *and* $C(\boldsymbol{A}^T) \cap N(\boldsymbol{A}) = \{\mathbf{0}\}$, *the row space is orthogonal to the null space, and they are orthogonal complements of one another,*

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

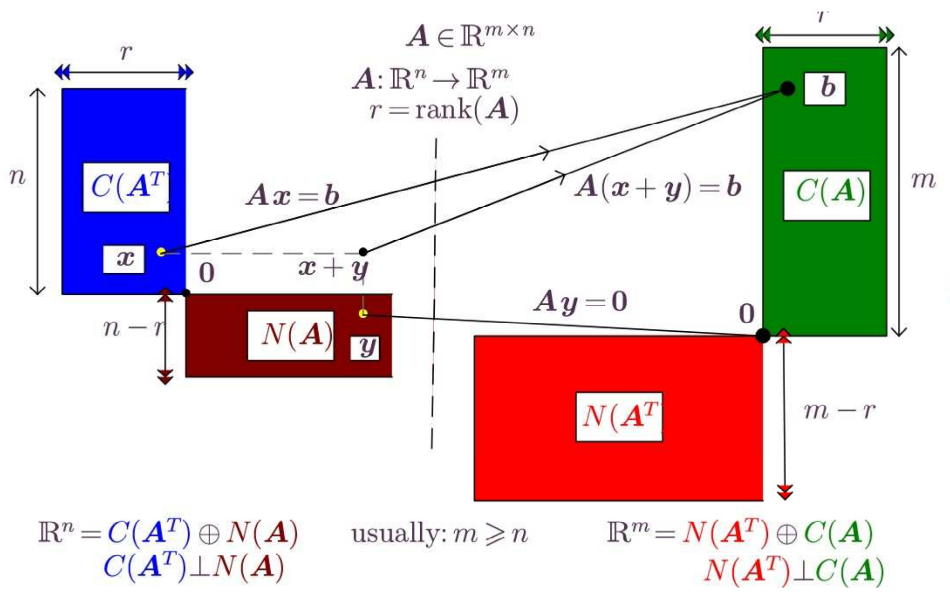

Figure 1. Graphical representation of the Fundamental Theorem of Linear Algebra, Gil Strang, *Amer. Math. Monthly* **100**, 848-855, 1993.

Consideration of equality between sets arises in proving the above theorem. A standard technique to show set equality $A = B$, is by double inclusion, $A \subseteq B \wedge B \subseteq A \Longrightarrow A = B$. This is shown for the statements giving the decomposition of the codomain $\mathbb{R}^m$. A similar approach can be used to decomposition of $\mathbb{R}^n$.

i. $C(\boldsymbol{A}) \perp N(\boldsymbol{A}^T)$ (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 $\boldsymbol{u}^T\boldsymbol{v} = (\boldsymbol{A}\boldsymbol{x})^T\boldsymbol{v} = \boldsymbol{x}^T\boldsymbol{A}^T\boldsymbol{v} = \boldsymbol{x}^T(\boldsymbol{A}^T\boldsymbol{v}) = \boldsymbol{x}^T\boldsymbol{0} = 0$, hence $\boldsymbol{u} \perp \boldsymbol{v}$ for arbitrary $\boldsymbol{u}, \boldsymbol{v}$, and $C(\boldsymbol{A}) \perp N(\boldsymbol{A}^T)$. □

ii. $C(\boldsymbol{A}) \cap N(\boldsymbol{A}^T) = \{\boldsymbol{0}\}$ ($\boldsymbol{0}$ 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 $\boldsymbol{x} = 0 \Longrightarrow \boldsymbol{b} = 0$, contradicting assumptions. Multiply equality $\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ on left by $\boldsymbol{x}^T$,

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

   thereby obtaining $\boldsymbol{b} = 0$, using norm property 3. Contradiction.

   □

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

   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 1 it results that $C(\boldsymbol{A}) \oplus N(\boldsymbol{A}^T) = \mathbb{R}^m$. □

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 $C(\boldsymbol{A}^T) \oplus N(\boldsymbol{A}) = \mathbb{R}^n$, etc.

3. The main problems of linear algebra

The FTLA asserts that the framework of linear algebra is complete, we can answer any question about linear combinations. What are the types of questions that are asked? These are known as the main problems of linear algebra.

1. *Least squares problem*. Given the components of vector $\boldsymbol{b} \in \mathbb{R}^m$ in the standard basis vectors, find the linear combination of the vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{R}^m$ that is “as close as possible' to $\boldsymbol{b}$. Assess discrepancy between $\boldsymbol{b}$ and a linear combination $\boldsymbol{A}\boldsymbol{x}$ through the two-norm

$$\min_{\boldsymbol{x} \in \mathbb{R}^n} \|\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}\|_2$$

2. *Solving a linear system*. Given the components of vector $\boldsymbol{b} \in \mathbb{R}^m$ in the standard basis vectors, what are the coordinates with respect to another set of vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{R}^m$?

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

   with $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ a matrix with column vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{R}^m$. Very often the matrix $\boldsymbol{A}$ is square $m = n$.

3. *Eigenproblem*. Given a square matrix $\boldsymbol{A} \in \mathbb{C}^{m \times m}$ are there linear combinations that leave the direction of a matrix-vector product unchanged?

$$\text{Given } \boldsymbol{A} \in \mathbb{C}^{m \times m}, \text{find } \boldsymbol{x} \in \mathbb{C}^m, \lambda \in \mathbb{C} \text{ such that } \boldsymbol{A}\boldsymbol{x} = \lambda \boldsymbol{x}.$$

   Note that the problem is now specified to allow complex components of $\boldsymbol{A}$. The linear algebra framework developed for vectors in $\mathbb{R}^m$ carries over with minimal modification to $\mathbb{C}^m$, and the eigenproblem requires consideration of complex values.