LECTURE 6: FUNDAMENTAL THEOREM OF LINEAR ALGEBRA

1. Vector Subspaces

1.1. Definitions

A central interest in scientific computation is to seek simple descriptions of complex objects. A typical situation is specifying an instance of some object of interest through an m-tuple $\boldsymbol{v} \in \mathbb{R}^m$ with large m. Assuming that addition and scaling of such objects can cogently be defined, a vector space is obtained, say over the field of reals with an Euclidean distance, E_m. Examples include for instance recordings of medical data (electroencephalograms, electrocardiograms), sound recordings, or images, for which m can easily reach into the millions. A natural question to ask is whether all the m real numbers are actually needed to describe the observed objects, or perhaps there is some intrinsic description that requires a much smaller number of descriptive parameters, that still preserves the useful idea of linear combination. The mathematical transcription of this idea is a vector subspace.

DEFINITION. (VECTOR SUBSPACE). *$\mathcal{U} = (U, S, +, \cdot)$, $U \neq \emptyset$, is a vector subspace of vector space $\mathcal{V} = (V, S, +, \cdot)$ over the same field of scalars S, denoted by $\mathcal{U} \leq \mathcal{V}$, if $U \subseteq V$ and $\forall a, b \in S$, $\forall \boldsymbol{u}, \boldsymbol{v} \in U$, the linear combination $a\boldsymbol{u} + b\boldsymbol{v} \in U$.*

The above states a vector subspace must be closed under linear combination, and have the same vector addition and scaling operations as the enclosing vector space. The simplest vector subspace of a vector space is the null subspace that only contains the null element, $U = \{\boldsymbol{0}\}$. In fact any subspace must contain the null element $\boldsymbol{0}$, or otherwise closure would not be verified for the particular linear combination $\boldsymbol{u} + (-\boldsymbol{u}) = \boldsymbol{0}$. If $U \subset V$, then $\mathcal{U}$ is said to be a *proper subspace* of $\mathcal{V}$, denoted by $\mathcal{U} < \mathcal{V}$.

- Setting $n - m$ components equal to zero in the real space $\mathcal{R}_m$ defines a proper subspace whose elements can be placed into a one-to-one correspondence with the vectors within $\mathcal{R}_n$. For example, setting component m of $\boldsymbol{x} \in \mathbb{R}^m$ equal to zero gives $\boldsymbol{x} = [\, x_1 \;\; x_2 \;\; \cdots \;\; x_{m-1} \;\; 0 \,]^T$ that while not a member of $\mathbb{R}^{m-1}$, it is in a one-to-one relation with $\boldsymbol{x}' = [\, x_1 \;\; x_2 \;\; \cdots \;\; x_{m-1} \,]^T \in \mathbb{R}^{m-1}$. Dropping the last component of $\boldsymbol{y} \in \mathbb{R}^m$, $\boldsymbol{y} = [\, y_1 \;\; y_2 \;\; \cdots \;\; y_{m-1} \;\; y_m \,]^T$ gives vector $\boldsymbol{y}' = [\, y_1 \;\; y_2 \;\; \cdots \;\; y_{m-1} \,] \in \mathbb{R}^{m-1}$, but this is no longer a one-to-one correspondence since for some given $\boldsymbol{y}'$, the last component y_m could take any value.

Vector subspaces arise in decomposition or partitioning of a vector space. The converse, composition of vector spaces $\mathcal{U} = (U, S, +, \cdot)$, $\mathcal{V} = (V, S, +, \cdot)$ is defined in terms of linear combination. A vector $\boldsymbol{x} \in \mathbb{R}^3$ can be obtained as the linear combination

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ x_2 \\ x_3 \end{bmatrix},$$

but also as

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 - a \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ a \\ x_3 \end{bmatrix},$$

for some arbitrary $a \in \mathbb{R}$. In the first case, $\boldsymbol{x}$ is obtained as a unique linear combination of a vector from the set $U = \{[\, x_1 \;\; 0 \;\; 0 \,]^T | \, x_1 \in \mathbb{R}\}$ with a vector from $V = \{[\, 0 \;\; x_2 \;\; x_3 \,]^T | \, x_2, x_3 \in \mathbb{R}\}$. In the second case, there is an infinity of linear combinations of a vector from V with another from $W = \{[\, x_1 \;\; x_2 \;\; 0 \,]^T | \, x_1, x_2 \in \mathbb{R}\}$ to the vector $\boldsymbol{x}$. This is captured by a pair of definitions to describe vector space composition.

DEFINITION. *Given two vector subspaces $\mathcal{U} = (U, S, +, \cdot)$, $\mathcal{V} = (V, S, +, \cdot)$ of the space $\mathcal{W} = (W, S, +, \cdot)$, the sum is the vector space $\mathcal{U} + \mathcal{V} = (U + V, S, +, \cdot)$, where the sum of the two sets of vectors U, V is $U + V = \{\boldsymbol{u} + \boldsymbol{v} | \, \boldsymbol{u} \in U, \boldsymbol{v} \in V\}$.*

DEFINITION. *Given two vector subspaces $\mathcal{U} = (U, S, +, \cdot)$, $\mathcal{V} = (V, S, +, \cdot)$ of the space $\mathcal{W} = (W, S, +, \cdot)$, the direct sum is the vector space $\mathcal{U} \oplus \mathcal{V} = (U \oplus V, S, +, \cdot)$, where the direct sum of the two sets of vectors U, V is $U \oplus V = \{\boldsymbol{u} + \boldsymbol{v} | \, \exists! \boldsymbol{u} \in U, \exists! \boldsymbol{v} \in V\}$. (unique decomposition)*

- Since the same scalar field, vector addition, and scaling is used , it is more convenient to refer to vector space sums simply by the sum of the vector sets $U + V$, or $U \oplus V$, instead of specifying the full tuplet for each space. This shall be adopted henceforth to simplify the notation.

In the previous example, the essential difference between the two ways to express $\boldsymbol{x} \in \mathbb{R}^3$ is that $U \cap V = \{\boldsymbol{0}\}$, but $V \cap W = \{[\ 0\ \ a\ \ 0\ ]^T | a \in \mathbb{R}\} \neq \{\boldsymbol{0}\}$, and in general if the zero vector is the only common element of two vector spaces then the sum of the vector spaces becomes a direct sum. In practice, the most important procedure to construct direct sums or check when an intersection of two vector subspaces reduces to the zero vector is through an inner product.

DEFINITION. *Two vector subspaces U, V of the real vector space $\mathbb{R}^m$ are orthogonal, denoted as $U \perp V$ if $\boldsymbol{u}^T \boldsymbol{v} = 0$ for any $\boldsymbol{u} \in U, \boldsymbol{v} \in V$.*

DEFINITION. *Two vector subspaces U, V of $U + V$ are orthogonal complements, denoted $U = V^\perp$, $V = U^\perp$ if they are orthogonal subspaces, $U \perp V$, and $U \cap V = \{\boldsymbol{0}\}$, i.e., the null vector is the only common element of both subspaces.*

The above concept of orthogonality can be extended to other vector subspaces, such as spaces of functions. It can also be extended to other choices of an inner product, in which case the term conjugate vector spaces is sometimes used. The concepts of sum and direct sum of vector spaces used linear combinations of the form $\boldsymbol{u} + \boldsymbol{v}$. This notion can be extended to arbitrary linear combinations.

DEFINITION. *In vector space $\mathcal{V} = (V, S, +, \cdot)$, the span of vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n \in V$, is the set of vectors reachable by linear combination*

$$\text{span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n\} = \{\boldsymbol{b} \in V | \ \exists x_1, \ldots, x_n \in S \text{ such that } \boldsymbol{b} = x_1 \boldsymbol{a}_1 + \ldots + x_n \boldsymbol{a}_n\}.$$

Note that for real vector spaces a member of the span of the vectors $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n\}$ is the vector $\boldsymbol{b}$ obtained from the matrix vector multiplication

$$\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x} = [\ \boldsymbol{a}_1\ \ \boldsymbol{a}_2\ \ \ldots\ \ \boldsymbol{a}_n\ ] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

From the above, the span is a subset of the co-domain of the linear mapping $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}$.

1.2. Vector subspaces of a linear mapping

The wide-ranging utility of linear algebra results from a complete characterization of the behavior of a linear mapping between vector spaces $\boldsymbol{f}: U \to V$, $\boldsymbol{f}(a\boldsymbol{u} + b\boldsymbol{v}) = a\boldsymbol{f}(\boldsymbol{u}) + b\boldsymbol{f}(\boldsymbol{v})$. For some given linear mapping the questions that arise are:

1. Can any vector within V be obtained by evaluation of $\boldsymbol{f}$?
2. Is there a single way that a vector within V can be obtained by evaluation of $\boldsymbol{f}$?

Linear mappings between real vector spaces $\boldsymbol{f}: \mathbb{R}^n \to \mathbb{R}^m$, have been seen to be completely specified by a matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$. It is common to frame the above questions about the behavior of the linear mapping $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}$ through sets associated with the matrix $\boldsymbol{A}$. To frame an answer to the first question, a set of reachable vectors is first defined.

DEFINITION. *The column space (or range) of matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ is the set of vectors reachable by linear combination of the matrix column vectors*

$$C(\boldsymbol{A}) = \text{range}(\boldsymbol{A}) = \{\boldsymbol{b} \in \mathbb{R}^m | \ \exists \boldsymbol{x} \in \mathbb{R}^n \text{ such that } \boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}\}.$$

By definition, the column space is included in the co-domain of the function $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}$, $C(\boldsymbol{A}) \subseteq \mathbb{R}^m$, and is readily seen to be a vector subspace of $\mathbb{R}^m$. The question that arises is whether the column space is the entire co-domain $C(\boldsymbol{A}) = \mathbb{R}^m$ that would signify that any vector can be reached by linear combination. If this is not the case then the column space would be a proper subset, $C(\boldsymbol{A}) \subset \mathbb{R}^m$, and the question is to determine what part of the co-domain cannot be reached by linear combination of columns of $\boldsymbol{A}$. Consider the orthogonal complement of $C(\boldsymbol{A})$ defined as the set vectors orthogonal to all of the column vectors of $\boldsymbol{A}$, expressed through inner products as

$$\boldsymbol{a}_1^T \boldsymbol{y} = 0, \boldsymbol{a}_2^T \boldsymbol{y} = 0, \ldots, \boldsymbol{a}_n^T \boldsymbol{y} = 0.$$

This can be expressed more concisely through the transpose operation

$$\boldsymbol{A} = [\ \boldsymbol{a}_1\ \ \boldsymbol{a}_2\ \ \cdots\ \ \boldsymbol{a}_n\ ], \boldsymbol{A}^T\boldsymbol{y} = \begin{bmatrix} \boldsymbol{a}_1^T \\ \boldsymbol{a}_2^T \\ \vdots \\ \boldsymbol{a}_n^T \end{bmatrix} \boldsymbol{y} = \begin{bmatrix} \boldsymbol{a}_1^T\boldsymbol{y} \\ \boldsymbol{a}_2^T\boldsymbol{y} \\ \vdots \\ \boldsymbol{a}_n^T\boldsymbol{y} \end{bmatrix},$$

and leads to the definition of a set of vectors for which $\boldsymbol{A}^T\boldsymbol{y} = \boldsymbol{0}$

DEFINITION. *The left null space (or cokernel) of a matrix $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ is the set*

$$N(\boldsymbol{A}^T) = \text{null}(\boldsymbol{A}^T) = \{\boldsymbol{y} \in \mathbb{R}^m | \boldsymbol{A}^T\boldsymbol{y} = \boldsymbol{0}\}.$$

Note that the left null space is also a vector subspace of the co-domain of $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}$, $N(\boldsymbol{A}^T) \subseteq \mathbb{R}^m$. The above definitions suggest that both the matrix and its transpose play a role in characterizing the behavior of the linear mapping $\boldsymbol{f} = \boldsymbol{A}\boldsymbol{x}$, so analagous sets are define for the transpose $\boldsymbol{A}^T$.

DEFINITION. *The row space (or corange) of a matrix $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ is the set*

$$R(\boldsymbol{A}) = C(\boldsymbol{A}^T) = \text{range}(\boldsymbol{A}^T) = \{\boldsymbol{c} \in \mathbb{R}^n | \exists \boldsymbol{y} \in \mathbb{R}^m\ \boldsymbol{c} = \boldsymbol{A}^T\boldsymbol{y}\} \subseteq \mathbb{R}^n$$

DEFINITION. *The null space of a matrix $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ is the set*

$$N(\boldsymbol{A}) = \text{null}(\boldsymbol{A}) = \{\boldsymbol{x} \in \mathbb{R}^n | \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}\} \subseteq \mathbb{R}^n$$

Examples. Consider a linear mapping $\boldsymbol{f}: \mathbb{R}^n \to \mathbb{R}^m$, defined by $\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x} = [\ y_1\ \ \cdots\ \ y_n\ ]^T$, with $\boldsymbol{A} \in \mathbb{R}^{m\times n}$.

1. For $n = 1$, $m = 3$,
$$\boldsymbol{A} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \boldsymbol{A}^T = [\ 1\ \ 0\ \ 0\ ],$$
the column space $C(\boldsymbol{A})$ is the y_1-axis, and the left null space $N(\boldsymbol{A}^T)$ is the y_2y_3-plane.
2. For $n = 2$, $m = 3$,
$$\boldsymbol{A} = \begin{bmatrix} 1 & -1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} = [\ \boldsymbol{a}_1\ \ \boldsymbol{a}_2\ ], \boldsymbol{A}^T = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix},$$
the columns of $\boldsymbol{A}$ are colinear, $\boldsymbol{a}_2 = -\boldsymbol{a}_1$, and the column space $C(\boldsymbol{A})$ is the y_1-axis, and the left null space $N(\boldsymbol{A}^T)$ is the y_2y_3-plane, as before.
3. For $n = 2$, $m = 3$,
$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \boldsymbol{A}^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix},$$
the column space $C(\boldsymbol{A})$ is the y_1y_2-plane, and the left null space $N(\boldsymbol{A}^T)$ is the y_3-axis.
4. For $n = 2$, $m = 3$,
$$\boldsymbol{A} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 0 & 0 \end{bmatrix}, \boldsymbol{A}^T = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \end{bmatrix},$$
the same $C(\boldsymbol{A})$, $N(\boldsymbol{A}^T)$ are obtained.
5. For $n = 3$, $m = 3$,
$$\boldsymbol{A} = \begin{bmatrix} 1 & 1 & 3 \\ 1 & -1 & -1 \\ 1 & 1 & 3 \end{bmatrix} = [\ \boldsymbol{a}_1\ \ \boldsymbol{a}_2\ \ \boldsymbol{a}_3\ ],$$
$$\boldsymbol{A}^T = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 3 & -1 & 3 \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_1^T \\ \boldsymbol{a}_2^T \\ \boldsymbol{a}_3^T \end{bmatrix}, \boldsymbol{A}^T\boldsymbol{y} = \begin{bmatrix} \boldsymbol{a}_1^T\boldsymbol{y} \\ \boldsymbol{a}_2^T\boldsymbol{y} \\ \boldsymbol{a}_3^T\boldsymbol{y} \end{bmatrix}$$

since $\boldsymbol{a}_3 = \boldsymbol{a}_1 + 2\boldsymbol{a}_2$, the orthogonality condition $\boldsymbol{A}^T\boldsymbol{y} = \boldsymbol{0}$ is satisfied by vectors of form $\boldsymbol{y} = [\ a\ \ 0\ \ -a\ ]$, $a \in \mathbb{R}$.

The above low dimensional examples are useful to gain initial insight into the significance of the spaces $C(\boldsymbol{A}), N(\boldsymbol{A}^T)$. Further appreciation can be gained by applying the same concepts to processing of images. A gray-scale image of size p_x by p_y pixels can be represented as a vector with $m = p_x p_y$ components, $\boldsymbol{b} \in [0,1]^m \subset \mathbb{R}^m$. Even for a small image with $p_x = p_y = 128 = 2^7$ pixels along each direction, the vector $\boldsymbol{b}$ would have $m = 2^{14}$ components. An image can be specified as a linear combination of the columns of the identity matrix

$$\boldsymbol{b} = \boldsymbol{I}\boldsymbol{b} = [\ \boldsymbol{e}_1\ \ \boldsymbol{e}_2\ \ \ldots\ \ \boldsymbol{e}_m\ ] \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix},$$

with b_i the gray-level intensity in pixel i. Similar to the inclined plane example from §1, an alternative description as a linear combination of another set of vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$ might be more relevant. One choice of greater utility for image processing mimics the behavior of the set $\{1, \cos t, \cos 2t, \ldots, \sin t, \sin 2t, \ldots\}$ that extends the second example in §1, would be for $m = 4$

$$\boldsymbol{A} = [\ \boldsymbol{a}_1\ \ \boldsymbol{a}_2\ \ \boldsymbol{a}_3\ \ \boldsymbol{a}_4\ ] = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

2. Linear dependence, vector space basis and dimension

2.1. Linear dependence

For the simple scalar mapping $f: \mathbb{R} \to \mathbb{R}$, $f(x) = ax$, the condition $f(x) = 0$ implies either that $a = 0$ or $x = 0$. Note that $a = 0$ can be understood as defining a zero mapping $f(x) = 0$. Linear mappings between vector spaces, $\boldsymbol{f}: U \to V$, can exhibit different behavior, and the condtion $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$, might be satisfied for both $\boldsymbol{x} \neq \boldsymbol{0}$, and $\boldsymbol{A} \neq \boldsymbol{0}$. Analogous to the scalar case, $\boldsymbol{A} = \boldsymbol{0}$ can be understood as defining a zero mapping, $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0}$.

In vector space $\mathscr{V} = (V, S, +, \cdot)$, vectors $\boldsymbol{u}, \boldsymbol{v} \in V$ related by a scaling operation, $\boldsymbol{v} = a\boldsymbol{u}$, $a \in S$, are said to be colinear, and are considered to contain redundant data. This can be restated as $\boldsymbol{v} \in \text{span}\{\boldsymbol{u}\}$, from which it results that $\text{span}\{\boldsymbol{u}\} = \text{span}\{\boldsymbol{u}, \boldsymbol{v}\}$. Colinearity can be expressed only in terms of vector scaling, but other types of redundancy arise when also considering vector addition as expressed by the span of a vector set. Assuming that $\boldsymbol{v} \notin \text{span}\{\boldsymbol{u}\}$, then the strict inclusion relation $\text{span}\{\boldsymbol{u}\} \subset \text{span}\{\boldsymbol{u}, \boldsymbol{v}\}$ holds. This strict inclusion expressed in terms of set concepts can be transcribed into an algebraic condition.

DEFINITION. *The vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n \in V$, are linearly dependent if there exist n scalars, $x_1, \ldots, x_n \in S$, at least one of which is different from zero such that*

$$x_1\boldsymbol{a}_1 + \ldots + x_n\boldsymbol{a}_n = \boldsymbol{0}.$$

Introducing a matrix representation of the vectors

$$\boldsymbol{A} = [\ \boldsymbol{a}_1\ \ \boldsymbol{a}_2\ \ \ldots\ \ \boldsymbol{a}_n\ ]; \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

allows restating linear dependence as the existence of a non-zero vector, $\exists \boldsymbol{x} \neq \boldsymbol{0}$, such that $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$. Linear dependence can also be written as $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0} \nRightarrow \boldsymbol{x} = \boldsymbol{0}$, or that one cannot deduce from the fact that the linear mapping $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}$ attains a zero value that the argument itself is zero. The converse of this statement would be that the only way to ensure $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ is for $\boldsymbol{x} = \boldsymbol{0}$, or $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0} \Rightarrow \boldsymbol{x} = \boldsymbol{0}$, leading to the concept of linear independence.

DEFINITION. *The vectors* $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n \in V$, *are linearly independent if the only n scalars,* $x_1, \ldots, x_n \in S$, *that satisfy*

$$x_1\boldsymbol{a}_1 + \ldots + x_n\boldsymbol{a}_n = \boldsymbol{0}, \tag{4}$$

are $x_1 = 0,\ x_2 = 0, \ldots, x_n = 0$.

2.2. Basis and dimension

Vector spaces are closed under linear combination, and the span of a vector set $\mathcal{B} = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots\}$ defines a vector subspace. If the entire set of vectors can be obtained by a spanning set, $V = \operatorname{span}\mathcal{B}$, extending $\mathcal{B}$ by an additional element $\mathcal{C} = \mathcal{B} \cup \{\boldsymbol{b}\}$ would be redundant since $\operatorname{span}\mathcal{B} = \operatorname{span}\mathcal{C}$. This is recognized by the concept of a basis, and also allows leads to a characterization of the size of a vector space by the cardinality of a basis set.

DEFINITION. *A set of vectors* $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n \in V$ *is a basis for vector space* $\mathcal{V} = (V, S, +, \cdot)$ *if*

1. $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ *are linearly independent;*
2. $\operatorname{span}\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\} = V$.

DEFINITION. *The number of vectors* $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n \in V$ *within a basis is the dimension of the vector space* $\mathcal{V} = (V, S, +, \cdot)$.

2.3. Dimension of matrix spaces

The domain and co-domain of the linear mapping $\boldsymbol{f}: U \to V, \boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}$, are decomposed by the spaces associated with the matrix $\boldsymbol{A}$. When $U = \mathbb{R}^n$, $V = \mathbb{R}^m$, the following vector subspaces associated with the matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ have been defined:

- $C(\boldsymbol{A})$ the column space of $\boldsymbol{A}$
- $C(\boldsymbol{A}^T)$ the row space of $\boldsymbol{A}$
- $N(\boldsymbol{A})$ the null space of $\boldsymbol{A}$
- $N(\boldsymbol{A}^T)$ the left null space of $\boldsymbol{A}$, or null space of $\boldsymbol{A}^T$

DEFINITION. *The rank of a matrix* $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ *is the dimension of its column space and is equal to the dimension of its row space.*

DEFINITION. *The nullity of a matrix* $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ *is the dimension of its null space.*

3. The FTLA

3.1. Partition of linear mapping domain and codomain

A partition of a set S has been introduced as 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 $\boldsymbol{0}$ belongs to only one member of the partition.

Linear mappings between vector spaces $\boldsymbol{f}: U \to V$ 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, $\boldsymbol{f}: \mathbb{R}^n \to \mathbb{R}^m$, in which case $\boldsymbol{A} \in \mathbb{R}^{m \times n}$,

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

◦ **Example 1.** Rotation by θ in $\mathbb{R}^2$ is obtained from

$$\boldsymbol{f}(\boldsymbol{e}_1)=\begin{bmatrix}\cos\theta\\ \sin\theta\end{bmatrix}, f(\boldsymbol{e}_2)=\begin{bmatrix}-\sin\theta\\ \cos\theta\end{bmatrix}$$

leading to

$$\boldsymbol{A}=\begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}.$$

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 $\boldsymbol{A}$,

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

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 2. *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. $\mathcal{U}\cap\mathcal{V}=\{\boldsymbol{0}\}$.

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 $\mathcal{W}=\mathcal{U}\oplus\mathcal{V}\Rightarrow\mathcal{U}\cap\mathcal{V}=\{\boldsymbol{0}\}$, consider $\boldsymbol{w}\in\mathcal{U}\cap\mathcal{V}$. Since $\boldsymbol{w}\in\mathcal{U}$ and $\boldsymbol{w}\in\mathcal{V}$ write*

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

and since expression $\boldsymbol{w}=\boldsymbol{u}+\boldsymbol{v}$ is unique, it results that $\boldsymbol{w}=\boldsymbol{0}$. 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 $\boldsymbol{x}=\boldsymbol{0}$, 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 if $U\perp V$, and $U\cap V=\{\boldsymbol{0}\}$. When $U\leq\mathbb{R}^m$, the orthogonal complement of U is denoted as $U^\perp$, $U\oplus U^\perp=\mathbb{R}^m$.

3.2. FTLA statement

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

1. *$C(\boldsymbol{A})\oplus N(\boldsymbol{A}^T)=\mathbb{R}^m$, the direct sum of the column space and left null space is the codomain of the mapping*
2. *$C(\boldsymbol{A}^T)\oplus N(\boldsymbol{A})=\mathbb{R}^n$, the direct sum of the row space and null space is the domain of the mapping*
3. *$C(\boldsymbol{A})\perp N(\boldsymbol{A}^T)$ and $C(\boldsymbol{A})\cap N(\boldsymbol{A}^T)=\{\boldsymbol{0}\}$, 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 $C(\boldsymbol{A}^T)\cap N(\boldsymbol{A})=\{\boldsymbol{0}\}$, the row space is orthogonal to the null space, and they are orthogonal complements of one another,*

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

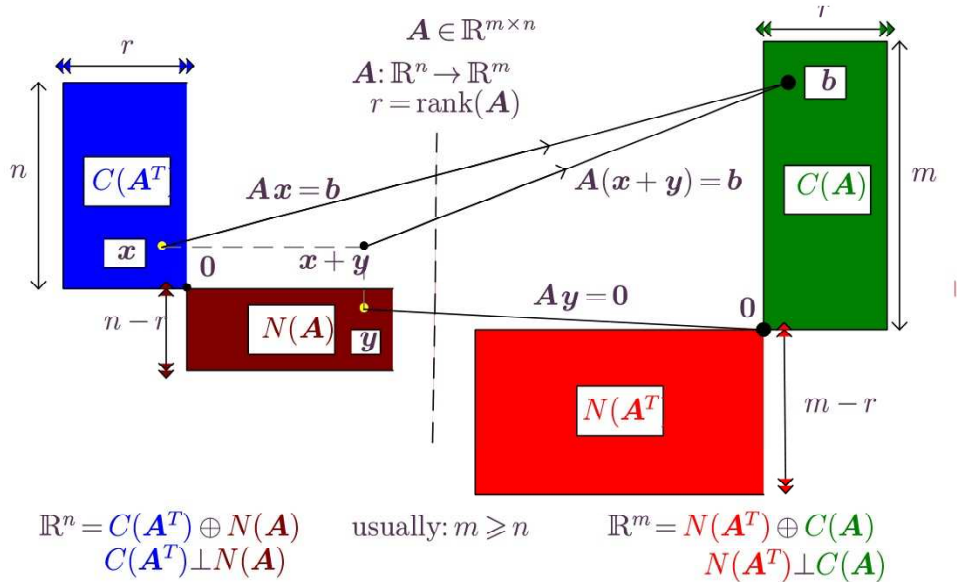

Figure 1. Graphical represenation of the Fundamental Theorem of Linear Algebra, Gil Strang, *Amer. Math. Monthly* **100**, 848-855, 1993.

Consideration of equality between sets arises in proving the above theorem. A standard technique to show set equality $A = B$, is by double inclusion, $A \subseteq B \wedge B \subseteq A \Rightarrow A = B$. 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 $\boldsymbol{x} = 0 \Rightarrow \boldsymbol{b} = 0$, contradicting assumptions. Multiply equality $\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ on left by $\boldsymbol{x}^T$,

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

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

Summary.

- Vector subspaces are subsets of a vector space closed under linear combination
- The simplest vector subspace is $\{\mathbf{0}\}$
- Linear mappings are represented by matrices
- Associated with matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ that represents mapping $f: \mathbb{R}^n \to \mathbb{R}^m$ are four fundamental subspaces:
    1. $C(\boldsymbol{A}) \leqslant \mathbb{R}^m$ the column space of $\boldsymbol{A}$ containing vectors $\boldsymbol{b}$ reachable by $\boldsymbol{A}\boldsymbol{x}, \boldsymbol{b} = \boldsymbol{A}x$
    2. $N(\boldsymbol{A}^T) \leqslant \mathbb{R}^m$ the left null space of $\boldsymbol{A}$ containing vectors $\boldsymbol{y}$ orthogonal to columns $\boldsymbol{A}$, $\boldsymbol{A}^T\boldsymbol{y} = 0$
    3. $C(\boldsymbol{A}^T) \leqslant \mathbb{R}^n$ the row space of $\boldsymbol{A}$
    4. $N(\boldsymbol{A}) \leqslant \mathbb{R}^n$ the null space of $\boldsymbol{A}$