MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Department of Electrical Engineering and Computer Science

**6.241: Dynamic Systems—Fall 2003**

RECITATION 1

# Linear Algebra Review

This handout comprises some of the Linear Algebra concepts that you need for this class. Use it as guide on your review. Not all the topics will be covered in the recitation, but we encourage you to work through them on your own. If you would like to see complete proofs of the results, some can be found in the 6.241 Lectures notes, the rest is the bulk of many standard books on the subject. A few references can be found at the end of this recitation note.

## 1 Outline

You have probably studied $\mathbb{R}^2$ and its algebraic structure, such as vector addition and multiplication by a scalar (real number). Similarly, you must have seen the inner product and the notion of Euclidean distance. These geometrical properties of $\mathbb{R}^2$ bring powerful and intuitive notions that created the basis for many achievements in the most diverse fields of Science and Engineering. The study of Vector Spaces started once the scientific community realized that those *nice properties* were not restricted to $\mathbb{R}^2$, or even $\mathbb{R}^3$. They exist in other cases, provided that we have the appropriate algebraic structure, which we call linear or vector space structure.

In this note, we will study Vector spaces and their properties in the following order:

- Vector Spaces.
  - Linear independence
  - (Hammel) Basis and dimension
  - Examples
  - Subspaces
  - Examples
- Inner Product Spaces (Vector space with inner product)
  - Orthogonal complement subspace
  - Examples

In this sequence, we move to linear operators according to the following:

- Linear Operators between finite dimensional vector spaces
  - Representation in Matrix Form
  - Examples

Finally, we devote some attention to the properties of a general matrix $\mathbf{A}$:

- Column space and range of $\mathbf{A}$
- Row space and the kernel (Null space) of $\mathbf{A}$
- Relations between the kernel of $\mathbf{A}$ and the range of $\mathbf{A}'$, the Hermitian conjugate of $A$.

As we go along the study of Vector spaces, try to visualize all the concepts in $\mathbb{R}^2$, or $\mathbb{R}^3$ if you prefer. We suggest you don't try it for $\mathbb{R}^n$, with $n \geq 4$, unless you want to torture yourself.

# 2 Vector Spaces and Inner Product Spaces

## 2.1 Vector Spaces

The following is the definition of a Vector Space:

**Definition 2.1** *Let $\mathbb{F}$ be a field given by $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$. A non-empty set $\mathbb{V}$ together with the binary operations "+"(vector addition) and "."(scalar multiplication) is a* **Vector Space** *provided that all the following are satisfied:*

- **Closedness under addition:**
$$\vec{v}_1, \vec{v}_2 \in \mathbb{V} \implies \vec{v}_1 + \vec{v}_2 \in \mathbb{V} \tag{1}$$
- **Zero Element:** $\exists \vec{0} \in \mathbb{V}$ *such that*
$$\vec{v} \in \mathbb{V} \implies \vec{v} + \vec{0} = \vec{v} \tag{2}$$
- **Commutativity under addition**
$$\vec{v}_1, \vec{v}_2 \in \mathbb{V} \implies \vec{v}_1 + \vec{v}_2 = \vec{v}_2 + \vec{v}_1 \tag{3}$$
- **Associativity under addition**
$$\vec{v}_1, \vec{v}_2, \vec{v}_3 \in \mathbb{V} \implies (\vec{v}_1 + \vec{v}_2) + \vec{v}_3 = \vec{v}_1 + (\vec{v}_2 + \vec{v}_3) \tag{4}$$
- **Closedness under scalar multiplication**
$$\alpha \in \mathbb{F}, \vec{v} \in \mathbb{V} \implies \alpha.\vec{v} \in \mathbb{V} \tag{5}$$
- **Distributivity**
$$\alpha_1, \alpha_2 \in \mathbb{F}, \vec{v} \in \mathbb{V} \implies (\alpha_1 + \alpha_2).\vec{v} = \alpha_1.\vec{v} + \alpha_2.\vec{v} \tag{6}$$
$$\alpha \in \mathbb{F}, \vec{v}_1, \vec{v}_2 \in \mathbb{V} \implies \alpha.(\vec{v}_1 + \vec{v}_2) = \alpha.\vec{v}_1 + \alpha.\vec{v}_2 \tag{7}$$
- **Associativity under scalar multiplication**
$$\alpha_1, \alpha_2 \in \mathbb{F}, \vec{v} \in \mathbb{V} \implies (\alpha_1\alpha_2).\vec{v} = \alpha_1.(\alpha_2.\vec{v}) \tag{8}$$
- **Multiplication by unity:**
$$\vec{v} \in \mathbb{V} \implies 1.\vec{v} = \vec{v} \tag{9}$$

### 2.1.1 Examples of Vector Spaces

The following are examples and counter-examples of Vector spaces. Check the veracity of the examples yourself. We will use these same examples as we go through the main properties of Vector spaces. The examples marked with (A) are more advanced, which means that it might be a little harder to check.

**Example 2.1** *The Euclidean space $\mathbb{R}^n$ is a vector space over the field $\mathbb{R}$. Note that sometimes the field of a vector space is implicitly specified. Since we only consider the real ($\mathbb{R}$) or the complex ($\mathbb{C}$) fields, you have to check which of the two, if any, makes sense. You can check that $\mathbb{R}^n$ is not a vector space over $\mathbb{C}$, as such we just have to write $\mathbb{R}^n$ and the field is implicitly $\mathbb{R}$.*

**Example 2.2** *On the other hand, $\mathbb{C}^n$ may be associated with two different vector spaces, one for the real field ($\mathbb{R}$) and another for $\mathbb{C}$, the complex field. Notice that even if you have the same set of elements, changing the field gives a different vector space. If we do not specify the field, we use the convention that $\mathbb{C}^n$ is the vector space over the complex field $\mathbb{C}$.*

**Example 2.3** *The positive cone $\mathbb{R}^n_+$ is not a vector space:*

$$\mathbb{R}^n_+ = \{(x_1, \ldots, x_n) | x_i \in \mathbb{R}, x_i > 0\} \tag{10}$$

**Example 2.4** *The vector space of real sequences:*

$$\mathbb{R}^\infty = \{(\ldots, x_{-1}, x_0, x_1, \ldots) | x_i \in \mathbb{R}\} \tag{11}$$

**Example 2.5** *The vector space of complex sequences (we convention that it is over the complex field $\mathbb{C}$):*

$$\mathbb{C}^\infty = \{(\ldots, x_{-1}, x_0, x_1, \ldots) | x_i \in \mathbb{C}\} \tag{12}$$

**Example 2.6** *The space of polynomials of order up to $n$, with real coefficients:*

$$\mathbb{P}^n = \{p(t) = a_0 + a_1 t + \ldots + a_n t^n | a_i \in \mathbb{R}\} \tag{13}$$

**Example 2.7** *(A) The space of real functions $f(t)$ with a finite number of discontinuities:*

$$\mathbb{CO}^0 = \{f : \mathbb{R} \to \mathbb{R} |\ f \text{ is not continuous on a finite number of points}\} \tag{14}$$

**Example 2.8** *(A) The space of real functions $f(t)$ such that its $n$th derivative has a finite number of discontinuities:*

$$\mathbb{CO}^n = \{f : \mathbb{R} \to \mathbb{R} |\ f^n \text{ is not continuous on a finite number of points}\} \tag{15}$$

### 2.1.2 Linearly independent vectors

Given a vector space $\mathbb{V}$ over a field $\mathbb{F}$, a finite linear combination $\vec{v}$ is just a weighted sum vectors:

$$\vec{v} = \sum_{i=0}^{q} \alpha_i \vec{v}_i,\ \vec{v}_i \in \mathbb{V}, \alpha_i \in \mathbb{F} \tag{16}$$

**Definition 2.2 Linear Independence** *A set of vectors $\{\vec{v}_1, \vec{v}_2 \ldots\}$, not necessarily finite, is linearly independent if one vector cannot be obtained as a finite linear combination of the others. Equivalently $\mathbb{I} = \{\vec{v}_1, \vec{v}_2 \ldots\}$, with $\vec{0} \notin \mathbb{I}$, is linearly independent if and only if:*

$$\alpha_i \in \mathbb{F}, \vec{v}_i \in \mathbb{I},\ \sum_{i=0}^{q} \alpha_i \vec{v}_i = \vec{0} \implies \alpha_i = 0 \tag{17}$$

*If the set is not linearly independent, we say it is linearly dependent.*

**Example 2.9** *In* $\mathbb{R}^2$ *the set* $\{(0,1),(1,0)\}$ *is linearly independent, but* $\{(1,2),(2,4)\}$ *is linearly dependent. You can also try to convince yourself that in* $\mathbb{R}^2$ *every 3 vectors are always linearly dependent. Also,* $\mathbb{R}^n$ *has, at most,* $n$ *linearly independent vectors.*

For a given set of vectors $\{\vec{v}_1, \ldots, \vec{v}_q\}$ of a vector space $\mathbb{V}$ over a field $\mathbb{F}$, we also define the set of all possible finite linear combinations, the span:

$$span(\{\vec{v}_1, \ldots, \vec{v}_q\}) = \{\bar{v} = \sum_{i=0}^{q} \alpha_i \vec{v}_i | \alpha_i \in \mathbb{F}\} \tag{18}$$

The span of an infinite set of vectors $\mathbb{W} = \{\vec{w}_1, \vec{w}_2, \ldots\}$ is given by:

$$span(\mathbb{W}) = \{\bar{v} = \sum_{i=0}^{q} \alpha_i \vec{w}_i | \vec{w}_i \in \mathbb{W}, \alpha_i \in \mathbb{F}, q \in \mathbb{N}_+\} \tag{19}$$

**Example 2.10** *Again, in* $\mathbb{R}^2$*, the* $span(\{(0,1)\})$ *is the "vertical axis" and* $span(\{(1,0),(0,1)\}) = \mathbb{R}^2$.

### 2.1.3 (Hamel) Basis and Dimension

The basis of a vector space $\mathbb{V}$ over a field $\mathbb{F}$ is a set of vectors $\mathbb{B}$, not necessarily finite, satisfying:

- $\mathbb{B}$ is linearly independent
- $span(\mathbb{B}) = \mathbb{V}$

As you study the properties of vector spaces, you will realize that the basis of a vector space is an extremely important concept. The following are properties of a basis of a given vector space $\mathbb{V}$ over a field $\mathbb{F}$:

- Every element $\vec{v} \in \mathbb{V}$ can be uniquely expressed as a finite sum of the form:

  $$\vec{v} = \sum_{i=1}^{q} \alpha_i \vec{w}_i \tag{20}$$

  where $\vec{w}_i \in \mathbb{B}$.
- The basis of a vector space is not unique, in fact, every vector space has an infinite number of basis.
- If $\mathbb{B}_1$ and $\mathbb{B}_2$ are basis of $\mathbb{V}$ over $\mathbb{F}$, then they have the same number of elements. That number is an intrinsic characteristic of the vector space which we designate as dimension. In general we write:

  $$dim(\mathbb{V}) = \text{number of elements of } \mathbb{B}_1 = \text{number of elements of } \mathbb{B}_2 \tag{21}$$
- A basis is the smallest set of vectors satisfying $span(\mathbb{B}) = \mathbb{V}$.
- $\vec{0}$ is never an element of the basis.

**Example 2.11** *Since* $span(\{(1,0),(0,1)\}) = \mathbb{R}^2$*, the set* $\mathbb{B} = \{(1,0),(0,1)\}$ *is a basis of* $\mathbb{R}^2$*. It also shows that the dimension of* $\mathbb{R}^2$ *is 2.*

**Example 2.12** *The set of canonical vectors* $\{\vec{e}_1, \ldots, \vec{e}_n\}$ *is a basis for* $\mathbb{R}^n$*, where a canonical vector is described as:*

$$\vec{e}_i = (x_1, \ldots, x_n),\ x_j = \begin{cases} 1 & \textit{if i=j} \\ 0 & \textit{Otherwise} \end{cases} \tag{22}$$

*The same set is also a basis for* $\mathbb{C}^n$*. The dimension of these spaces is* $n$*.*

**Example 2.13** *Consider the vector space of all real sequences with a finite number of nonzero terms:*

$$\mathbb{R}_f^\infty = \{(\ldots, x_{-1}, x_0, x_1, \ldots) \in \mathbb{R}^\infty | \exists m \in \mathbb{N}, \textit{such that } x_i = 0 \textit{ for } |i| > m\} \tag{23}$$

*The set* $\mathbb{B} = \{\vec{\delta}_i | i \in \mathbb{N}\}$ *is a basis for* $\mathbb{R}_f^\infty$*, where the following defines a canonical sequence:*

$$\vec{\delta}_i = (\ldots, x_{-1}, x_0, x_1, \ldots),\ x_j = \begin{cases} 1 & \textit{if i=j} \\ 0 & \textit{Otherwise} \end{cases} \tag{24}$$

*The space* $\mathbb{R}_f^\infty$ *is infinite dimensional.*

**Example 2.14** *The set* $\mathbb{B} = \{1, t, t^2, \ldots, t^n\}$ *is a basis for* $\mathbb{P}^n$*. Consequently,* $dim(\mathbb{P}^n) = n + 1$*.*

### 2.1.4 Subspaces

**Definition 2.3** *(Subspace) Consider a vector space* $\mathbb{V}$ *over the field* $\mathbb{F}$*. A subspace* $\mathbb{S}$ *of* $\mathbb{V}$ *is a vector space that satisfies the following:*

- $\mathbb{S}$ *is a vector space over the field* $\mathbb{F}$*.*
- $\mathbb{S} \subset \mathbb{V}$

*Equivalently, a subspace* $\mathbb{S}$ *is a nonempty subset of* $\mathbb{V}$ *that satisfies two conditions:*

1. $\alpha x \in \mathbb{S}$ *for all* $x \in \mathbb{S}$ *and any scalar* $\alpha$*, and*
2. $x + y \in \mathbb{S}$ *for all* $x,\ y \in \mathbb{S}$

*Note that the subspace must contain the* $\vec{0}$ *vector (set* $\alpha = 0$ *in rule 1).*

**Example 2.15** *Given* $m \leq n$*, check that* $\mathbb{R}^m$ *is a subspace of* $\mathbb{R}^n$ *as well as* $\mathbb{C}^m$ *is a subspace of* $\mathbb{C}^n$*.*

**Example 2.16** *If* $\mathbb{V}$ *is a vector space with base* $\mathbb{B}$*, then* $span(\mathbb{B}_S)$ *is a subspace of* $\mathbb{V}$ *provided that* $\mathbb{B}_S \subset \mathbb{B}$*. You can also show that if* $\mathbb{S}$ *is a subspace of a vector space* $\mathbb{V}$*, then* $dim(\mathbb{V}) \geq dim(\mathbb{S})$*.*

**Example 2.17** *The set* $\{(x, y, z) \in \mathbb{R}^3 | x + 2y + 3z = u\}$ *is a subspace of* $\mathbb{R}^3$ *if and only if* $u = 0$*. If* $u = 0$*, what is the dimension of this subspace? (Hint: check that* $\mathbb{B} = \{(-2, 1, 0), (-3, 0, 1)\}$ *is a basis for the space of solutions of* $x + 2y + 3z = 0$*.)*

**Example 2.18** *Any bounded subset of* $\mathbb{R}^n$ *is not a subspace.*

**Example 2.19** *(A) For* $n \geq m$*,* $\mathbb{CO}^m$ *is a subspace of* $\mathbb{CO}^n$*.*

## 2.2 Inner Product Spaces

Again, we generalize the notions of angle and projection of $\mathbb{R}^2$.

**Definition 2.4** *A vector space $\mathbb{V}$ over a field $\mathbb{F}$ is an inner product space provided that there exists a binary operator $< \cdot, \cdot >: \mathbb{V} \times \mathbb{V} \rightarrow \mathbb{C}$ that satisfies the following conditions:*

- *If $\vec{v}_1, \vec{v}_2 \in \mathbb{V}$, then $< \vec{v}_1, \vec{v}_2 >= \overline{< \vec{v}_2, \vec{v}_1 >}$*
- *If $\vec{v}_1, \vec{v}_2, \vec{v}_3 \in \mathbb{V}$, then $< \vec{v}_1 + \vec{v}_2, \vec{v}_3 >=< \vec{v}_1, \vec{v}_3 > + < \vec{v}_2, \vec{v}_3 >$*
- *If $\alpha \in \mathbb{F}, \vec{v}_1, \vec{v}_2 \in \mathbb{V}$, then $< \alpha\vec{v}_1, \vec{v}_2 >= \alpha < \vec{v}_1, \vec{v}_2 >$*
- *If $\vec{v} \in \mathbb{V}$, then $< \vec{v}, \vec{v} >\geq 0$.*
- *If $\vec{v} \in \mathbb{V}$, then we define a distance or norm $\|\vec{v}\|$ as:*

$$\|\vec{v}\|^2 =< \vec{v}, \vec{v} > \tag{25}$$

- $\|\vec{v}\| = 0 \iff \vec{v} = \vec{0}$

  You can check (not easy) that the norm defined above satisfies:

- **Cauchy Schwartz:** $| < \vec{v}_1, \vec{v}_2 > | \leq \|\vec{v}_1\|\|\vec{v}_2\|$
- **Triangle inequality** $\|\vec{v}_1 + \vec{v}_2\| \leq \|\vec{v}_1\| + \|\vec{v}_2\|$

**Definition 2.5** *A basis $\mathbb{B}$ of an inner-product space $\mathbb{V}$ is an orthogonal basis, provided that:*

$$\forall \vec{v}_1 \neq \vec{v}_2, \vec{v}_1, \vec{v}_2 \in \mathbb{B}, < \vec{v}_1, \vec{v}_2 >= 0 \tag{26}$$

**Example 2.20** *Clearly $\mathbb{R}^n$ is an inner-product space where, the Euclidean inner-product is given by:*

$$< (x_1, \ldots, x_n), (y_1, \ldots, y_n) >= \sum_{i=1}^{n} x_i y_i \tag{27}$$

*Similarly $\mathbb{C}^n$ has an inner-product given by*

$$< (x_1, \ldots, x_n), (y_1, \ldots, y_n) >= \sum_{i=1}^{n} \overline{x_i} y_i \tag{28}$$

*In any of these cases, the canonical basis of example 2.12 is also an orthogonal basis. Given an hermitian and positive definite matrix $Q \in \mathbb{C}^{n\times n}$, if we write the coordinates of the two vectors $\vec{x}, \vec{y} \in \mathbb{V}$, with $\mathbb{V} = \mathbb{R}^n$ or $\mathbb{V} = \mathbb{C}^n$, in the following column matrix form:*

$$\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \vec{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \tag{29}$$

*then $< \vec{x}, \vec{y} >= \vec{x}' Q \vec{y}$ is also valid inner product. Note that if $Q$ is the identity then we obtain the Euclidean inner-product.*

**Example 2.21** *(A) The space of $l_2$ sequences given by:*

$$l_2 = \{(\ldots, x_{-1}, x_0, x_1, \ldots) \in \mathbb{R}^\infty : \sum_{-\infty}^{\infty} x_i^2 < \infty\} \tag{30}$$

*with inner product given by*

$$< (\ldots, x_{-1}, x_0, x_1, \ldots), (\ldots, y_{-1}, y_0, y_1, \ldots) >= \sum_{i=-\infty}^{\infty} x_i y_i \tag{31}$$

*You can also show that $l_2$ is a subspace of $\mathbb{R}^\infty$.*

**Example 2.22** *(A) Motivated by the last example, you can also show that the following is an inner product space:*

$$L_2 = \{f(t) \in \mathbb{CO}^0 : \int_{-\infty}^{\infty} f(t)^2 < \infty\} \tag{32}$$

*where the inner product is defined by $< f, g >= \int_{-\infty}^{\infty} f(t)g(t)$. Also, $L_2$ is a subspace of $\mathbb{CO}^0$.*

**Example 2.23** *(A) The same space with different inner products generates distinct inner-product spaces. The space $\mathbb{P}^n$, for instance, admits the following two inner products:*

$$\text{(first inner product)} < p_1, p_2 >_1 = \sum_{i=1}^{n} a_i b_i \tag{33}$$

$$\text{( second inner product)} < p_1, p_2 >_2 = \int_{-1}^{1} p_1(t) p_2(t) \tag{34}$$

*where $p_1(t) = a_0 + \ldots + a_n t^n$ and $p_2(t) = b_0 + \ldots + b_n t^n$. Notice that, using the first inner product, the set $\{1, t, \ldots, t^n\}$ is an orthogonal basis. If we adopt the second inner-product, we conclude that an orthogonal basis is given by the "famous" Legendre polynomials $\{1, t, \frac{1}{2}(3t^2 - 1), \ldots\}$. They can be computed using a process called the Gram-Schmidt orthogonalization. Such process allows the construction of an orthogonal basis for any inner product space.*

**Definition 2.6** *If $\mathbb{S}$ is a subset of an inner product space $\mathbb{V}$, then $\mathbb{S}^\perp$ is the orthogonal subspace defined as:*

$$\mathbb{S}^\perp = \{\vec{v} \in \mathbb{V} | \forall \vec{s} \in \mathbb{S}, < \vec{v}, \vec{s} >= 0\} \tag{35}$$

**Example 2.24** *In $\mathbb{R}^2$, we have $\{(1,1)\}^\perp = span(\{(-1,1)\})$*

# 3 Linear Operators

Now, we move to the study of linear operators. A linear operator is a "function" between vector spaces. A simple example is the rotation around the origin of vectors, by a fixed angle, in $\mathbb{R}^2$. Another simple example is the multiplication by a scalar.

**Definition 3.1 (Linear Operator)** *A linear operator $\mathcal{L} : \mathbb{V}^I \to \mathbb{V}^O$ between two vector spaces $\mathbb{V}^I$ and $\mathbb{V}^O$ over the same field $\mathbb{F}$, must satisfy:*

$$\mathcal{L}(\alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2) = \alpha_1 \mathcal{L}(\vec{v}_1) + \alpha_2 \mathcal{L}(\vec{v}_2) \tag{36}$$

*where $\vec{v}_1, \vec{v}_2 \in \mathbb{V}^I$ and $\alpha_1, \alpha_2 \in \mathbb{F}$.*

The following are properties of linear operators:

- Given two vector spaces $\mathbb{V}^I$ and $\mathbb{V}^O$ over the same field $\mathbb{F}$, the set of all possible linear operators is a vector space over $\mathbb{F}$.
- If $\mathbb{B}^I = \{\vec{w}_1^I, \vec{w}_2^I, \ldots\}$ is a basis for $\mathbb{V}^I$ and $\vec{v} \in \mathbb{V}^I$ is an arbitrary vector, then for some $q$:

$$\mathcal{L}(\vec{v}) = \mathcal{L}(\sum_{i=1}^{q} \alpha_i \vec{w}_i^I) = \sum_{i=1}^{q} \alpha_i \mathcal{L}(\vec{w}_i^I) \tag{37}$$

  where $\vec{v} = \sum_{i=1}^{q} \alpha_i \vec{w}_i^I$ is the unique representation of $\vec{v}$ in terms of the basis vectors.

The above property is very important, because it shows that we only really need to know $\mathcal{L}(\cdot)$ on the elements of the basis $\mathbb{B}^I$ to completely characterize the operator over the whole space $\mathbb{V}^I$. Equivalently, it shows that all the vectors that $\mathcal{L}$ will "produce" are some linear combination of $\mathcal{L}(\vec{w}_i^I)$, where $\vec{w}_i^I$ are elements of $\mathbb{B}^I$. The following concept addresses this property:

**Definition 3.2** *(**Range and rank**) The range of a linear operator $\mathcal{L} : \mathbb{V}^I \to \mathbb{V}^O$ is defined as:*

$$\mathcal{R}(\mathcal{L}) = span(\{\mathcal{L}(\vec{w}_1^I), \mathcal{L}(\vec{w}_2^I), \ldots\}) \tag{38}$$

*where $\mathbb{B}^I = \{\vec{w}_1^I, \vec{w}_2^I, \ldots\}$ is a basis for $\mathbb{V}^I$. As such, the range is a subspace of $\mathbb{V}^O$ that can be generated by action of $\mathcal{L}$ over $\mathbb{V}^I$. In rigorous terms:*

$$\vec{z} \in \mathcal{R}(\mathcal{L}) \Longleftrightarrow \exists \vec{v} \in \mathbb{V}^I \text{ such that } \mathcal{L}(\vec{v}) = \vec{z} \tag{39}$$

*The rank of a linear operator is the dimension of its range $rank(\mathcal{L}) = dim(\mathcal{R}(\mathcal{L}))$.*

**Definition 3.3** *(**Null Space**) The null space of a linear operator is given by: (show that it is a subspace of $\mathbb{V}^I$)*

$$\mathcal{N}(\mathcal{L}) = \{\vec{v} \in \mathbb{V}^I : \mathcal{L}(\vec{v}) = \vec{0}\} \tag{40}$$

## 3.1 Linear Operators between finite dimensional vector spaces

Now we show that if $\mathbb{V}^I$ and $\mathbb{V}^O$ are finite dimensional, then any linear operator $\mathcal{L} : \mathbb{V}^I \to \mathbb{V}^O$ can be represented by a matrix.

Consider a linear operator $\mathcal{L} : \mathbb{V}^I \to \mathbb{V}^O$ between two finite dimensional vector spaces $\mathbb{V}^I$ and $\mathbb{V}^O$ over the same field $\mathbb{F}$. Let $\mathbb{B}^I = \{\vec{w}_1^I, \ldots, \vec{w}_n^I\}$ and $\mathbb{B}^O = \{\vec{w}_1^O, \ldots, \vec{w}_n^O\}$ be the basis of $\mathbb{V}^I$ and $\mathbb{V}^O$. Define $a_{i,j} \in \mathbb{F}$ as the expansion that satisfies:

$$\sum_{i=1}^{m} a_{i,j} \vec{w}_i^O = \mathcal{L}(\vec{w}_j^I) \tag{41}$$

Notice that there always exist such expansion. Indeed $\vec{z} = \mathcal{L}(\vec{w}_j^I)$ is an element of $\mathbb{V}^O$ and, as such, it can be written as a linear combination of the elements of the basis $\mathbb{B}^O$. Now, for an arbitrary element $\vec{v} = \sum_{j=1}^{n} \alpha_j^I \vec{w}_j^I$, you can show that:

$$\mathcal{L}(\vec{v}) = \sum_{j=1}^{n} \alpha_j^I \mathcal{L}(\vec{w}_j^I) = \sum_{i=1}^{m} \alpha_i^O \vec{w}_i^O \tag{42}$$

where $\alpha_i^O$ is computed as:

$$\begin{bmatrix}\alpha_1^O\\ \vdots \\ \alpha_m^O\end{bmatrix} = \begin{bmatrix} a_{1,1} & \dots & a_{1,n}\\ \vdots & \ddots & \vdots \\ a_{m,1} & \dots & a_{m,n}\end{bmatrix}\begin{bmatrix}\alpha_1^I\\ \vdots \\ \alpha_n^I\end{bmatrix} \tag{43}$$

The analysis above shows that, given finite basis $\mathbb{B}^I$ and $\mathbb{B}^O$, the linear operator $\mathcal{L} : \mathbb{V}^I \to \mathbb{V}^O$ can be equivalently represented by a matrix. Notice that if we change our choice for $\mathbb{B}^I$ and $\mathbb{B}^O$, the matrix will also change.

**Example 3.1** *Let $\mathbb{V}^I = \mathbb{V}^O = \mathbb{P}^3$, then the derivative operator $\mathcal{L}(p) = \frac{dp(t)}{dt}$ can be represented, under the basis $\mathbb{B}^I = \mathbb{B}^O = \{1, t, t^2, t^3\}$ as:*

$$A = \begin{bmatrix} 0 & 1 & 0 & 0\\ 0 & 0 & 2 & 0\\ 0 & 0 & 0 & 3\\ 0 & 0 & 0 & 0\end{bmatrix} \tag{44}$$

**Definition 3.4** ***(Range and Rank of a matrix A)*** *If $A \in \mathbb{F}^{m\times n}$ is a matrix, with $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$, then the range of A is defined as:*

$$\mathcal{R}(A) = span(\{\begin{bmatrix} a_{1,1}\\ \vdots \\ a_{m,1}\end{bmatrix}, \dots, \begin{bmatrix} a_{1,n}\\ \vdots \\ a_{m,n}\end{bmatrix}\}) \tag{45}$$

*The range of $A$ is also designated as the column space of $A$, for obvious reasons. Similarly, the rank of $A$ is defined as $rank(A) = dim(\mathcal{R}(A))$. The rank is the number of linearly independent columns of $A$, which constitute a basis for $\mathcal{R}(A)$.*

**Definition 3.5** ***(Null space of a matrix A)*** *Similarly, we define:*

$$\mathcal{N}(A) = \{\vec{v} \in \mathbb{F}^n | A\vec{v} = \vec{0}\} \tag{46}$$

*You can also show that $\mathcal{N}(A) = \mathcal{R}(A')^{\perp}$.*

Let $\mathcal{L} : \mathbb{V}^I \to \mathbb{V}^O$ be a linear operator between two finite dimensional vector spaces $\mathbb{V}^I$ and $\mathbb{V}^O$ over the same field $\mathbb{F}$ with basis $\mathbb{B}^I$ and $\mathbb{B}^O$. If $A \in \mathbb{F}^{m\times n}$ is the matrix representation of $\mathcal{L}$, then we can compute its range and null space as:

$$\mathcal{R}(\mathcal{L}) = \{\sum_{i=1}^{m} \alpha_i^O \vec{w}_i^O \mid \vec{w}_i^O \in \mathbb{B}^O \text{ and } \begin{bmatrix}\alpha_1^O\\ \vdots \\ \alpha_m^O\end{bmatrix} \in \mathcal{R}(A)\} \tag{47}$$

$$\mathcal{N}(\mathcal{L}) = \{\sum_{i=1}^{n} \alpha_i^I \vec{w}_i^I \mid \vec{w}_i^I \in \mathbb{B}^I \text{ and } \begin{bmatrix}\alpha_1^I\\ \vdots \\ \alpha_n^I\end{bmatrix} \in \mathcal{N}(A)\} \tag{48}$$

**Example 3.2** *Consider $\mathcal{L} : \mathbb{R}^n \to \mathbb{R}^m$ and a matrix representation $A$. If we use the canonical basis, then $\mathcal{R}(A) = \mathcal{R}(\mathcal{L})$ and $\mathcal{N}(A) = \mathcal{N}(\mathcal{L})$.*

**Example 3.3** *Consider the linear operator $\frac{d}{dt} : \mathbb{P}^2 \to \mathbb{P}^2$ and basis $\mathbb{B}^I = \mathbb{B}^O = \{1, t, t^2\}$. Then $\mathcal{R}(\frac{d}{dt}) = span(\{1, t\})$ and $\mathcal{N}(\frac{d}{dt}) = span(\{1\})$.*

There exist a number of equalities involving the range and the null space of a matrix. You can look them up on the references bellow. Some of them will appear as part of your problem set $\#1$.

### 3.2 Linear Equations

Most importantly, these notions are used to characterize the possible types of linear equations involving a linear operator $\mathcal{L} : \mathbb{V}^I \to \mathbb{V}^O$ between two inner-product spaces $\mathbb{V}^I$ and $\mathbb{V}^O$. In the following, we assume that $\vec{y} \in \mathbb{V}^O$ is given and that we want to find $\vec{x} \in \mathbb{V}^I$ such that $\mathcal{L}(\vec{x}) = \vec{y}$:

- **Invertible** If $\mathcal{R}(\mathcal{L}) = \mathbb{V}^O$, then by definition $\mathcal{L}(\vec{x}) = \vec{y}$ has always a solution. If $dim(\mathcal{N}(\mathcal{L})) = 0$, then the solution is unique. When that happens the matrix representation is a square and invertible matrix $A$ and the solution is obtained by inverting it.
- **Under-constrained** If $\mathcal{R}(\mathcal{L}) = \mathbb{V}^O$, $dim(\mathcal{N}(\mathcal{L})) > 0$ and $\vec{x}$ is a solution, then you can check that $\vec{x} + \vec{z}, \vec{z} \in \mathcal{N}(\mathcal{L})$ is also a solution. That implies that the solution is not unique. Here we use least squares to choose the solution with smallest $\|\vec{x}\|$.
- **Over-constrained** If $\mathcal{R}(\mathcal{L}) \subsetneq \mathbb{V}^O$, then $\mathcal{L}(\vec{x}) = \vec{y}$ has a solution only if $\vec{y} \in \mathcal{R}(\mathcal{L})$. If $\vec{y} \notin \mathcal{R}(\mathcal{L})$, then the best we can do is to choose the solution that minimizes $\|\vec{y} - \mathcal{L}(\vec{x})\|$

The solution to the problems above will be discussed next week.

## References


[1] Kailath, T., Sayed, A., and Hassibi, B. “Linear Estimation.” *Prentice Hall*: 2000.

[2] Rudin, W. “Principles of Methematical Analysis.” *McGraw Hill*: 1964.

[3] Strang, G. “Linear Algebra and Its Applications,” *Academic Press, Inc.*: 1976.