ADVANCED LINEAR ALGEBRA

Adair Antonio da Silva Neto

September 6, 2022

Contents

Chapter 1

Vector Spaces

In this chapter, we'll proceed with an overview of elementary linear algebra, covering the definition of vector spaces, bases and coordinates, linear transformations and matrices, rank, nullity, inner product, normal and self-adjoint operators, and diagonalization. The proofs in this chapter will be skipped.

1.1 Vector Spaces

> Loosely speaking, linear algebra is that branch of mathematics which treats the common properties of algebraic systems which consist of a set, together with a reasonable notion of a 'linear combination' of elements in the set.

Definition 1.1.1 (Vector Space)**.** A **vector space** (or **linear space**) V over a field $\mathbb{F}$ is a set with a binary operation '+' on V (called **addition**) and an action '$\cdot$' of $\mathbb{F}$ on V (called **scalar multiplication**) such that, for any $x, y \in V$ and $a, b \in \mathbb{F}$, $x + y \in V$ (closed under addition) and $a \cdot x \in V$ (invariant under scalar multiplication) satisfying:

1. $x + y = y + x$.
2. $(x + y) + z = x + (y + z)$.
3. There exists $0 \in V$ such that $x + 0 = x$ for all $x \in V$.
4. For all $x \in V$, there exists $y \in V$ such that $x + y = 0$.
5. There exists $1 \in \mathbb{F}$ such that $1 \cdot x = x$ for all $x \in V$.
6. $a \cdot (b \cdot x) = (a \cdot b) \cdot x$.
7. $a \cdot (x + y) = a \cdot x + a \cdot y$.
8. $(a + b) \cdot x = a \cdot x + b \cdot x$.

We'll refer to the elements of V as **vectors** and to the elements of $\mathbb{F}$ as **scalars**.

In the following pages, we'll use V to denote a vector space and $\mathbb{F}$ to denote a field. And 'iff.' means 'if and only if'.

Example 1.1.1 (Some Vector Spaces).

1. The **zero-dimensional space**. The set $V = \{0\}$ under some field $\mathbb{F}$.
2. The **field $\mathbb{F}$ as a one-dimensional coordinate space**. A field (e.g. $\mathbb{C}$) can be interpreted as a vector space of a subfield of it (e.g. $\mathbb{R}$).
3. The **n-tuple space** $\mathbb{F}^n$.
4. The **space of $m \times n$ matrices** $\mathbb{F}^{m\times n}$.
5. **Function spaces** $F(S)$. Which maps S into the field $\mathbb{F}$.
6. The **space of polynomial functions over a field** $\mathbb{F}$.

Some immediate conclusions follow from this definition.

Lemma 1.1.1 (Basic Properties). For all $x \in V$ and $a \in \mathbb{F}$, the following properties hold:

1. $\underset{\in \mathbb{F}}{0} \cdot x = \underset{\in V}{0}$
2. $\underset{\in \mathbb{F}}{(-a)} \cdot x = \underset{\in V}{-(a \cdot x)} = \underset{\in V}{a \cdot (-x)}$
3. $a \cdot \underset{\in V}{0} = \underset{\in V}{0}$

The basic motivation of Linear Algebra is to solve systems of linear equations. The concept of linear combination is of essential character in solving these systems and inspires the definition of matrix multiplication and linear transformations.

Definition 1.1.2 (Linear combinations). Let $S \subseteq V$, $S \neq \varnothing$.

A vector $v \in V$ is a **linear combination** of S if it can be written as

$$v = a_1 u_1 + a_2 u_2 + \ldots + a_n u_n = \sum_{i=1}^{n} a_i u_i$$

for some vectors $u_1, \ldots, u_n \in S$ and scalars $a_1, \ldots, a_n \in \mathbb{F}$.

1.1.1 Subspaces

Definition 1.1.3 (Subspace). Let V be a vector space over a field $\mathbb{F}$. A subset $W \subseteq V$ is a **subspace** of V if W is itself a vector space with respect to the addition and scalar multiplication on V.

Theorem 1.1.2 (Criteria for Subspaces). Let $W \subseteq V$. Then W is a subspace of V iff.

1. $0 \in W$.
2. $x + y \in W$ for all $x, y \in W$ (closed under addition).
3. $c \cdot x \in W$ for all $c \in \mathbb{F}$ and $x \in W$ (closed under scalar multiplication).

However, we can simplify this check a little more.

Theorem 1.1.3 (New Criteria for Subspaces)**.** Let $W \subseteq V$. Then W is a subspace of V iff. for any $x, y \in W$ and $c \in \mathbb{F}$, we have that $cx + y \in W$.

The conditions that an arbitrary vector in V must satisfy in order to belong to W are called **linear conditions**. A combination of linear conditions is also a linear condition. In other words, we have the next theorem.

Theorem 1.1.4 (Intersection of subspaces is a subspace)**.** If $W_1, \ldots, W_n$ are subspaces of V, then $W = \bigcap_{i=1}^{n} W_i$ is also a subspace of V.

Definition 1.1.4 (Span)**.** Let $S \subseteq V$. The **subspace spanned** by S (or **span** of S), denoted by $\text{Span}(S)$, is the intersection of all subspaces of V which contain S.

We define the $\text{Span}(\varnothing) = \{0\}$.

The following theorem gives an equivalent definition.

Theorem 1.1.5 (Equivalent Definition for Span)**.** The **span** of S is the subset of V consisting of all linear combinations of S.

$$\text{Span}(S) = \{a_1 u_1 + \ldots + a_n u_n : n \in \mathbb{N}, a_i \in \mathbb{F}, u_i \in S\}$$

Theorem 1.1.6 (Properties of the Span)**.** Let S be any subset of V, not necessarily a subspace. Then,

1. $\text{Span}(S)$ is a subspace of V.
2. Any subspace of V containing S also must contain $\text{Span}(S)$.

Definition 1.1.5 (Generation of Spaces)**.** Let $S \subseteq V$. We say that S **generates** (or **spans**) V if $\text{Span}(S) = V$.

1.1.2 Bases and Dimension

Definition 1.1.6 (Linear Dependence)**.** A subset S of V is **linearly dependent** if there exists a finite number of distinct vectors $u_1, \ldots, u_n \in S$ and scalars $a_1, \ldots, a_n \in \mathbb{F}$, with at least one $a_i \neq 0$, such that

$$a_1 u_1 + \ldots + a_n u_n = 0$$

And $S \subseteq V$ is **linearly independent** if it is not linearly dependent, i.e., no non-trivial linear combination of $u_1, \ldots, u_n$ vanishes.

Theorem 1.1.7 (Criteria for Linear Dependence)**.** Let $S_1 \subseteq S_2 \subseteq V$.

1. If S_1 is linearly dependent, then S_2 is also linearly dependent.
2. If S_2 is linearly independent, then S_1 is also linearly independent.

3. Let $S \subseteq V$ be linearly independent, and $v \in V$ such that $v \notin S$. Then, $S \cup \{v\}$ is linearly dependent iff. $v \in \mathrm{Span}(S)$.

Definition 1.1.7 (Basis)**.** A **basis** for V is a subset of V which is both linearly independent and generates V.

Example 1.1.2. Let S be the subset of $\mathbb{F}^n$ containing

$$\begin{aligned} e_1 &= (1,0,0,\ldots,0) \\ e_2 &= (0,1,0,\ldots,0) \\ &\vdots \\ e_n &= (0,0,0,\ldots,1) \end{aligned}$$

Clearly, these vectors span $\mathbb{F}^n$ and are linearly independent. Then this set is a basis for $\mathbb{F}^n$ and is called the **standard basis** of $\mathbb{F}^n$.

An alternative characterization of vector spaces is given by the following theorem.

Theorem 1.1.8. A subset of vectors $\{u_1,\ldots,u_n\}$ of V is a basis iff. every $v \in V$ can be uniquely written in the form

$$v = a_1u_1 + \ldots + a_nu_n$$

for some $a_i \in \mathbb{F}$.

Theorem 1.1.9 (Replacement Theorem)**.** Let V be a vector space generated by $G \subseteq V$ with $|G| = n$, and L be a linearly independent subset of V, $|L| = m$. Then $m \leq n$, and there exists $H \subseteq G$ such that $|H| = n-m$ and $L \cup H$ generates V.

In other words, if V is a vector space spanned by a finite set of vectors $u_1,\ldots,u_n$, then any independent set of vectors in V is finite and contains no more than n elements.

The next theorem guarantees that every basis has the same cardinality, i.e., the number of elements in the basis does not depend on the basis.

Theorem 1.1.10. If V is a finitely generated vector space, then every basis of V has the same number of elements in it.

Definition 1.1.8 (Dimension)**.** If V is a finitely generated vector space, we define the **dimension** of V, denoted $\dim(V)$, as the cardinality of a basis for V.

Corollary 1.1.11. Let $n = \dim V < \infty$. Then

1. Any subset of V which contains more than n vectors is linearly dependent.
2. No subset of V which contains fewer than n vectors can span V.

Lemma 1.1.12. Let S be a linearly independent subset of a vector space V. If $v \in V$ is not in the subspace spanned by S, then the set obtained by adjoining v to S is linearly independent.

Theorem 1.1.13. If W is a subspace of a finite-dimensional vector space V, every linearly independent subset of W is finite and is a part of a finite basis for W.

A corollary of this theorem is that proper subspaces have smaller dimension.

Corollary 1.1.14 (Monotonicity of dimension)**.** Let W be a subspace of V with $\dim(V) < \infty$. Then

$$\dim(W) \leq \dim(V)$$

If the equality $\dim(W) = \dim(V)$ holds, then $V = W$.

Corollary 1.1.15 (Extension of a basis)**.** If $W = \{w_1, \ldots, w_m\}$ is a linearly independent set of vectors in a finite-dimensional vector space V, then there exists a basis of V that contains W.

Theorem 1.1.16. If W_1 and W_2 are both finite-dimensional subspaces of V, then $W_1 + W_2$ is finite-dimensional and

$$\dim W_1 + \dim W_2 = \dim(W_1 \cap W_2) + \dim(W_1 + W_2)$$

Definition 1.1.9 (Maximal)**.** Let $E = \{v_1, \ldots, v_n\}$ be a set of vectors in V and let $F = \{v_{i_1}, \ldots, v_{i_m}\}$ be a linearly independent subset of E. If every element in E can be expressed as a linear combination of the elements of F, then F is said to be **maximal**.

The number of elements in a maximal subset equals the dimension of the span of E and is called the **rank**.

Definition 1.1.10 (Flags)**.** A sequence of subspaces $V_0 \subset V_1 \subset \ldots \subset V_n$ of the space V is said to be a **flag**.

More generally, a sequence of subsets $S_0 \subset S_1 \subset \ldots \subset S_n$ is called **increasing filtering**.

A flag is said to be **maximal** if $V_0 = \{0\}$, $\bigcup V_i = V$ and there's no subspace between other two, i.e., if $V_i \subset M \subset V_{i+1}$ then either $V_i = M$ or $V_{i+1} = M$.

Notice that given any basis $\{u_1, \ldots, u_n\}$ of V, we can construct a flag by setting $V_0 = \{0\}$ and $V_i = \operatorname{span}(\{u_1, \ldots, u_i\})$ for $i \geq 1$.

Theorem 1.1.17. The dimension of a vector space V equals the length of any maximal flag of V.

The next theorem is an example of application of Zorn's lemma.

Theorem 1.1.18. Every vector space has a basis.

1.1.3 Coordinates

The coordinates of a vector relative to a basis will be the coefficients that are used to represent the vector as a linear combination of the vectors in the basis. For example, if $(v_1, \ldots v_n)$ is an arbitrary vector in $\mathbb{R}^n$ and $e_1, \ldots, e_n$ is the standard basis for $\mathbb{R}^n$, then we express

$$v = (v_1, \ldots, v_n) = \sum_{i=1}^{n} v_i e_i$$

However, for this expression to be adequately defined, the vectors in the basis must be ordered. To put it another way, we must look at our basis as a sequence instead of a set to distinguish its i-th element.

Definition 1.1.11 (Ordered Basis). Let $\dim(V) < \infty$. An **ordered basis** for V is a basis for V with a fixed order on its vectors.

With this definition, we say that v_i is the ith **coordinate of v relative to the ordered basis**. And we use $[v]_\beta$ to denote the coordinates of v concerning the ordered basis β. More precisely,

Definition 1.1.12 (Coordinates). Let $\beta = \{v_1, \ldots, v_n\}$ be an ordered basis for V. Then any vector $x \in V$ can be written uniquely as

$$x = a_1 v_1 + \ldots + a_n v_n$$

for $a_1, \ldots, a_n \in \mathbb{F}$.

We define the **coordinate vector** as

$$[x]_\beta = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \in \mathbb{F}^n$$

Now, what happens with the coordinates when we change from one basis to another?

Let $\beta = \{\beta_1, \ldots, \beta_n\}$ and $\gamma = \{\gamma_1, \ldots, \gamma_n\}$ be two ordered bases for the finite-dimensional space V. And notice that we can write every vector of the basis γ as a linear combination of the vectors of β as follows:

$$\begin{aligned}
\gamma_1 &= a_{11} \cdot \beta_1 + a_{21} \cdot \beta_2 + \ldots + a_{n1} \cdot \beta_n \\
\gamma_2 &= a_{12} \cdot \beta_1 + a_{22} \cdot \beta_2 + \ldots + a_{n2} \cdot \beta_n \\
&\vdots \\
\gamma_n &= a_{1n} \cdot \beta_1 + a_{2n} \cdot \beta_2 + \ldots + a_{nn} \cdot \beta_n
\end{aligned}$$

where each a_{ij} is a scalar.

Thus, for each $i \in \{1, 2, \ldots, n\}$, the coordinates vector of γ_i in the basis β is given by

$$[\gamma_i]_\beta = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{bmatrix}$$

With this algorithm, we obtain the coordinates of each vector in the basis γ concerning the basis β. And we form the **transition matrix**, also called **change-of-basis matrix**, from β to γ:

$$P_{\beta \to \gamma} = \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ a_{21} & \ldots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{nn} \end{bmatrix}$$

Note that each column is formed by the coordinates of $\gamma_1, \ldots, \gamma_n$ with respect to the basis β.

Theorem 1.1.19. Let V be an n-dimensional vector space and let $\beta = \{u_1, \dots, u_n\}$ and $\gamma = \{u'_1, \dots, u'_n\}$ be two ordered bases of V. Then there is a unique and invertible $n \times n$ matrix P such that

1. $[u]_\beta = P[u]_\gamma$,
2. $[u]_\gamma = P^{-1}[u]_\beta$,

for every vector $u \in V$. And the columns of P are given by

$$P_j = [u'_j]_\beta,\ j = 1, \dots, n$$

Example 1.1.3 (Change of basis)**.** Consider β the standard basis of $\mathbb{R}^3$ and

$$\gamma = \{(1,0,1), (1,1,1), (1,1,2)\}$$

Find the transition matrix $P_{\gamma \to \beta}$.

Solution: The first step is to write each vector of β as a linear combination of the vectors of γ. I.e.,

$$\begin{aligned}
(1,0,0) &= a_{11} \cdot (1,0,1) + a_{21} \cdot (1,1,1) + a_{31} \cdot (1,1,2) \\
&= 1 \cdot (1,0,1) + 1 \cdot (1,1,1) - 1 \cdot (1,1,2) \\
(0,1,0) &= a_{12} \cdot (1,0,1) + a_{22} \cdot (1,1,1) + a_{32} \cdot (1,1,2) \\
&= -1 \cdot (1,0,1) + 1 \cdot (1,1,1) + 0 \cdot (1,1,2) \\
(0,0,1) &= a_{13} \cdot (1,0,1) + a_{23} \cdot (1,1,1) + a_{33} \cdot (1,1,2) \\
&= 0 \cdot (1,0,1) - 1 \cdot (1,1,1) + 1(1,1,2)
\end{aligned}$$

With these values, we form the transition matrix:

$$P_{\gamma \to \beta} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & -1 \\ -1 & 0 & 1 \end{bmatrix}$$

1.1.4 The Row and Column Spaces of a Matrix

Before heading to next section, we introduce some useful nomenclature and results.

Definition 1.1.13 (Row Space)**.** Let A be an $m \times n$ matrix over the field $\mathbb{F}$. We define the **row space** as the subspace of $\mathbb{F}^n$ generated by the rows of A. The dimension of the row space is called **row rank**.

Theorem 1.1.20.

1. Row-equivalent matrices have the same row space.
2. The non-zero lines of a row-reduced echelon matrix form a basis for its row space.
3. If W is a subspace of $\mathbb{F}^n$ such that $\dim W \leq m$, then there exists a unique row-reduced

echelon matrix $m \times n$ over $\mathbb{F}$ whose row space is W.

4. Every matrix is row-equivalent to one, and only one, row reduced echelon matrix.

5. Two matrices are row-equivalent iff. they have the same row space.

1.2 Linear Transformations

In plain words, a linear transformation (or linear mapping) is a function from a vector space to another which preserves the structure of a vector space. More precisely,

1.2.1 Basic Definitions

Definition 1.2.1 (Linear Transformation). Let V and W be two vector spaces over the same field $\mathbb{F}$. A **linear transformation** $T : V \longrightarrow W$ is a function satisfying:

1. $T(x+y) = T(x) + T(y)$, for all $x, y \in V$.

2. $T(cx) = cT(x)$, for all $x \in V$, $c \in \mathbb{F}$.

Put it another way, a linear mapping is a **homomorphism** of additive groups.

Theorem 1.2.1 (Properties).

1. If T is a linear transformation, then $T(0) = 0$.

2. $T\left(\sum_{i=1}^{n} a_i x_i\right) = \sum_{i=1}^{n} a_i T(x_i)$ for all $x_i \in V$, $a_i \in \mathbb{F}$.

3. A function $T : V \longrightarrow W$ is a linear transformation iff. $T(cx+y) = cT(x) + T(y)$ for all $x, y \in V$, $c \in \mathbb{F}$.

Example 1.2.1. Let $\mathbb{F}$ be a field and V be the space of polynomial functions $f : \mathbb{F} \longrightarrow \mathbb{F}$ given by

$$f(x) = c_0 + c_1 x + \ldots + c_k x^k$$

Define

$$(Df)(x) = c_1 + 2c_2 x + \ldots + kc_k x^{k-1}$$

Then D is a linear transformation called the differentiation operator.

Example 1.2.2. Given the field of real numbers $\mathbb{R}$ and $V = \mathscr{C}(\mathbb{R})$, we define

$$T(f(x)) = \int_0^x f(t)\ dt$$

which is a linear transformation.

How can we define linear transformations? The easiest way is to define its values on a basis and then linearly extend it to the whole space. The next theorem says this process returns a well defined linear mapping.

Theorem 1.2.2. Let $\{v_1,...,v_n\}$ be a basis for V. Then for any vectors $w_1,...,w_n \in W$, there exists exactly one linear transformation $T : V \longrightarrow W$ such that

$$T(v_i) = w_i, \text{ for } 1 \leq i \leq n$$

Definition 1.2.2 (Null space and Range)**.** Let $T : V \longrightarrow W$ be a linear transformation.

1. The **null space** (or **kernel**) of T is

$$\ker(T) = \{x \in V : T(x) = 0\}$$

2. The **range** of T is the image V under T, i.e.,

$$\text{Im}(T) = \{y \in W : y = T(x), x \in V\}$$

The dimension of the range is called the **rank** of T and the dimension of the kernel is called the **nullity** of T.

Theorem 1.2.3. The null space of T ker(T) is a subspace of V and Im(T) is a subspace of W.

Theorem 1.2.4 (The Dimension Theorem (Rank–Nullity))**.** If $\dim(V) < \infty$, then

$$\dim(V) = \dim(\ker(T)) + \dim(\text{Im}(T))$$

i.e., $\dim(V) = \text{nullity}(T) + \text{rank}(T)$.

Definition 1.2.3 (Injection and Surjection)**.** Let $T : V \longrightarrow W$ be a linear transformation.

1. T is **injective** if $T(v) = T(u)$ implies $v = u$, for all $u, v \in V$.
2. T is **surjective** if for every $w \in W$ there exists $v \in V$ such that $T(v) = w$.
3. T is **bijective** if T is injective and surjective.

Theorem 1.2.5.

- T is injective iff. $\ker(T) = \{0\}$.
- T is surjective iff. $\text{Im}(T) = W$.

Theorem 1.2.6. Assume $\dim(V) = \dim(W)$. Then the following affirmations are equivalent:

1. T is injective.
2. T is surjective.
3. T is bijective.
4. $\dim(\text{Im}(T)) = \dim(V)$.

1.2.2 The Algebra of Linear Transformations

Theorem 1.2.7. Let $T, U : V \longrightarrow W$ be linear transformations. We define, for all $x \in V$ and $a \in \mathbb{F}$,

1. $(T+U)(x) = T(x) + U(x)$;
2. $(aT)(x) = aT(x)$.

Then $T+U$ and $a \cdot U$ are also linear transformations from V to W.

Theorem 1.2.8 (Space of Linear Transformations)**.** Let $\mathscr{L}(V, W)$ be the set of all linear transformations from V to W. Then $\mathscr{L}(V, W)$ is a vector space over the same field $\mathbb{F}$ with respect to the operations defined above.

An alternative notation is $\mathscr{L}(V, W) = \hom_{\mathbb{F}}(V, W)$. When $V = W$, we write $\mathscr{L}(V)$.

Theorem 1.2.9. If V is an n-dimensional vector space and W is an m-dimensional vector space, both over $\mathbb{F}$, then the space $\mathscr{L}(V, W)$ has dimension mn.

Definition 1.2.4 (Composition of Linear Transformations)**.** Let V, W, Z be vector spaces. Let $T : V \longrightarrow W$ and $U : W \longrightarrow Z$ be linear transformations. Their **composition** is the function $UT : V \longrightarrow Z$ defined by $(UT)(x) = U(T(x))$ for all $x \in V$.

Theorem 1.2.10 (Composition is also linear)**.** If T and U are both linear transformations, then their composition UT is a linear transformation.

Definition 1.2.5 (Linear Operator)**.** A **linear operator** is a linear transformation from a vector space to itself. It is also called an **endomorphism**. The set of linear operators on a vector space V is denoted by $\mathscr{L}(V)$ or $\mathrm{End}_{\mathbb{F}}(V)$.

Remark that if U and T are linear operators on V, then the composition $U \circ T$ is also a a linear operator on V. The space $\mathscr{L}(V)$ has a 'multiplication' defined on it by composition. The operator $T \circ U$ is also defined, but in general $UT \neq TU$, i.e., the **Lie bracket** $[U, T] = UT - TU \neq 0$.

Lemma 1.2.11. Let U, T_1 and T_2 be linear operators on the vector space V and $c \in \mathbb{F}$. The following affirmations hold.

1. $IU = UI = U$;
2. $U(T_1 + T_2) = UT_1 + UT_2$ and $(T_1 + T_2)U = T_1U + T_2U$;
3. $c(UT_1) = (cU)T_1 = U(cT_1)$.

As a matter of fact, the vector space $\mathscr{L}(V)$, together with the composition operation, is known as a **linear algebra with identity**.

Definition 1.2.6 (Invertibility)**.** Let V, W be vector spaces, and $T : V \longrightarrow W$ a linear transformation.

1. A linear transformation $U : W \longrightarrow V$ is the **inverse** of T if $UT = I_V$ and $TU = I_W$, where I denotes the identity matrix.

2. T is invertible if it has an inverse.

Theorem 1.2.12 (Characterization of Inverses). Let V, W be vector spaces, and $T : V \longrightarrow W$ a linear transformation.

1. If T is invertible, then its inverse is unique, denoted by T^{-1}.

2. T is invertible iff. T is a bijection.

Lemma 1.2.13. Let $T : V \longrightarrow W$ be an invertible linear transformation, and $\dim(V) < \infty$. Then $\dim(V) = \dim(W)$.

To check whether a transformation T is injective, notice that if T is linear, then $T(u-v) = T(u) - T(v)$. Therefore, $T(u) = T(v)$ iff. $T(u-v) = 0$.

Definition 1.2.7 (Non-singular Transformations). A linear mapping T is **non-singular** if $T(v) = 0$ implies $v = 0$, i.e., the null space of T is $\{0\}$.

Hence, T is injective iff. T is non-singular. more than that, non-singular linear transformations are those which preserve linear independence.

Theorem 1.2.14. Let $T : V \longrightarrow W$ be a linear mapping. Then T is non-singular if and only if T carries each linearly independent subset of V onto a linearly independent subset of W.

Theorem 1.2.15. Let V and W be finite-dimensional vector spaces such that $\dim V = \dim W$. If T is a linear mapping from V into W, the following are equivalent:

1. T is invertible;

2. T is non-singular;

3. T is surjective;

4. If $\{v_1, \ldots, v_n\}$ is a basis for V, then $\{T(v_1), \ldots, T(v_n)\}$ is a basis for W;

5. There is some basis for V such that $\{T(v_1), \ldots, T(v_n)\}$ is a basis for W.

The set of invertible linear operators on a given space, with the operation of composition, provides an example of a group.

Definition 1.2.8 (Group). A **group** consists of the following:

1. A set G;

2. A rule (or operation) $\odot$ which associates with each pair of elements $x, y \in G$ an element $x \odot y$ in G satisfying

   - Associativity: $x \odot (y \odot z) = (x \odot y) \odot z$, for all $x, y, z \in G$;
   - Identity: There is an element e in G such that $e \odot x = x \odot e = x$, for every x in G;

- Inverse: To each element $x \in G$ there corresponds an element x^{-1} in G such that $x \odot x^{-1} = x^{-1} \odot x = e$.

Example 1.2.3. The following are examples of groups.

- **General linear group** $GL(n)$, formed by the set of non-singular $n \times n$ matrices with the operation of function composition.
- **Permutation group** S_n, of permutations of sets of n elements.
- **Special linear group** $SL(n)$, of $n \times n$ matrices with determinant equal to one.
- **Orthogonal group** $O(n)$, of $n \times n$ matrices such that $AA^t = I$, which is the group of isometries of Euclidean space that preserve a fixed point.
- **Special Orthogonal group** $SO(n)$, consisting of orthogonal matrices whose determinant is equal to one.
- **Unitary group** $U(n)$ of all complex $n \times n$ matrices satisfying $AA^* = 1$, where $A^* = \bar{A}^t$.
- **Special unitary group** $SU(n)$ of unitary matrices with determinant one.

1.2.3 Isomorphisms

Definition 1.2.9 (Isomorsphism)**.** Bijective linear mappings $T \in \mathscr{L}(V, W)$ are said to be **isomorphisms**, and the spaces V and W are called **isomorphic** if there exists an isomorphism between them.

If $T \in \mathscr{L}(V)$ is an isomorphism, then T is said to be an **automorphism**.

Remark that isomophism is an equivalence relation in the family of vector spaces.

Theorem 1.2.16. Every n-dimensional vector space over a field $\mathbb{F}$ is isomorphic to $\mathbb{F}^n$.

To convince yourself that this claim is true, it is enough to map every vector to its coordinates in a given basis.

A more general result states that the dimension of a space completely determines the space up to isomorphism. To put it another way, every finite subspace $S \subseteq V$ has the same dimension as the range $T(S)$, i.e., isomorphisms preserve dimension.

Theorem 1.2.17. Two finite-dimensional spaces V and W are isomorphic iff. they have the same dimension.

If the isomorphism does not depend on arbitrary choices, such as the basis, then it is called a **canonical** or **natural isomorphism**. This will be made precise when the language of categories is introduced.

Finally, note that the isomorphisms from a space to itself form a group with respect to the operation of function composition, which is exactly the general linear group we saw earlier.

1.2.4 Matrix Representation

Definition 1.2.10 (Matrix Representation)**.** Let V, W be vector spaces with ordered basis $\beta = \{v_1, \dots, v_n\}$ and $\gamma = \{w_1, \dots, w_m\}$, respectively.

Let $T : V \longrightarrow W$ be a linear transformation. Then the **matrix representation** of T with respect to β and γ is defined as the matrix $[T]_{\beta,\gamma} \in \mathbb{M}_{m\times n}(\mathbb{F})$ given by

$$[T]_{\beta,\gamma} = \begin{pmatrix} | & | & & | \\ [T(v_1)]_\gamma & [T(v_2)]_\gamma & \dots & [T(v_n)]_\gamma \\ | & | & & | \end{pmatrix}$$

where $[T(v_i)]_\gamma$ are the coordinates of the vector $T(v_i) \in W$ with respect to the ordered basis γ.

If $V = W$ and $\beta = \gamma$, we write $[T]_\beta$.

Theorem 1.2.18. Assume V, W are finite dimensional vector spaces with ordered basis β and γ. Let $T, U : V \longrightarrow W$ be linear transformations. Then,

1. $U = T$ iff. $[U]_{\beta,\gamma} = [T]_{\beta,\gamma}$;
2. $[T+U]_{\beta,\gamma} = [T]_{\beta,\gamma} + [U]_{\beta,\gamma}$;
3. $[aT]_{\beta,\gamma} = a[T]_{\beta,\gamma}$, for all $a \in \mathbb{F}$.

Theorem 1.2.19. Let V, W, Z be vector spaces with ordered basis α, β, γ respectively. Let $T : V \longrightarrow W$ and $U : W \longrightarrow Z$ be linear transformations. Then

$$[UT]_{\alpha,\gamma} = [U]_{\beta,\gamma}[T]_{\alpha,\beta}$$

Corollary 1.2.20. Let V be a finite vector space with ordered basis β. Let $T, U \in \mathscr{L}(V)$. Then $[UT]_\beta = [U]_\beta[T]_\beta$.

Theorem 1.2.21. Let V, W be finite dimensional vector spaces with ordered basis β and γ. Let $T : V \longrightarrow W$ be a linear transformation. For all $u \in V$,

$$[T(u)]_\gamma = [T]_{\beta,\gamma}[u]_\beta$$

Definition 1.2.11 (Invertibility for a Matrix)**.** A matrix $A \in \mathbb{M}_{m\times n}(\mathbb{F})$ is **invertible** if there exists $B \in \mathbb{M}_{m\times n}(\mathbb{F})$ such that $AB = BA = I$.

Theorem 1.2.22. Let V, W be finite dimensional vector spaces with ordered bases β and γ respectively. Let $T : V \longrightarrow W$ be a linear transformation. Then T is invertible iff. $[T]_{\beta,\gamma}$ is invertible. Moreover,

$$[T^{-1}]_{\gamma,\beta} = ([T]_{\beta,\gamma})^{-1}$$

Theorem 1.2.23. Let V be a finite-dimensional vector space and let

$$\beta = \{v_1, \dots, v_n\} \text{ and } \gamma = \{w_1, \dots, w_n\}$$

be ordered basis for V. Suppose that $T \in \mathscr{L}(V)$. If P is the matrix with columns $P_j = [w_j]_\beta$ (i.e. the coordinates of the j-th vector on the basis β), then

$$[T]_\gamma = P^{-1}[T]_\beta P$$

Alternatively, if U is the invertible operator defined by $U[v_j] = w_j$, then

$$[T]_\gamma = [U]_\beta^{-1}[T]_\beta[U]_\beta$$

Definition 1.2.12 (Similar Matrices)**.** Let A and B be $n \times n$ matrices. We say that B is **similar** to A if there exists an invertible $n \times n$ matrix P such that

$$B = P^{-1}AP$$

1.2.5 Product and Quotient Spaces

1.2.6 Dual Space

A concept that will help us in the study of subspaces, linear equations, and coordinates is the following.

Definition 1.2.13 (Linear Functional and Dual Space)**.** A linear transformation from the vector space V to its scalar field $\mathbb{F}$ is called a **linear functional**.

The set of linear functionals is denoted by V^* and is called **dual space** of V. In other words, $V^* = \mathscr{L}(V, \mathbb{F})$.

Example 1.2.4. Let $(c_1, \dots, c_n) \in \mathbb{F}^n$ and define $f : \mathbb{F}^n \longrightarrow \mathbb{F}$ by

$$f(x_1, \dots, x_n) = c_1x_1 + \dots + c_nx_n$$

Then f is a linear functional on $\mathbb{F}^n$.

Example 1.2.5 (Trace)**.** If A is an $n \times n$ matrix, the **trace** of A is the scalar

$$\operatorname{tr} A = A_{11} + A_{22} + \dots + A_{nn}$$

Remark that the trace function is a linear functional on the matrix space $\mathbb{M}_n$.

Remark. Suppose V is finite-dimensional. Then the dimension of the dual space is equal to the dimension of the space.

$$\dim V^* = \dim V$$

Definition 1.2.14 (Dual basis)**.** If $\beta = \{v_1, \dots, v_n\}$ is a basis of V then the **dual basis** of β is the set $\beta^* = \{f_1, \dots, f_n\}$, where each f_i is the linear functional on V such that

$$f_i(v_j) = \delta_{ij}$$

Theorem 1.2.24. Let V be a finite-dimensional vector space. Then the dual basis of a basis of V is a basis of V^*.

Proof. Let $\beta = \{v_1, \ldots, v_n\}$ be a basis for V. Then there exists a unique linear functional f_i on V such that

$$f_i(v_j) = \delta_{ij}$$

for each i.

With this process, we obtain n distinct linear functionals $f_1, \ldots, f_n$ on V.

To show that $f_1, \ldots, f_n$ are linearly independent, suppose that $c_1, \ldots, c_n \in \mathbb{F}$ are such that

$$c_1 f_1 + \ldots + c_n f_n = 0$$

Since $(c_1 f_1 + \ldots + c_n f_n)(v_j) = c_j$ for each $j = 1, \ldots, n$, we know that $c_1 = \ldots = c_n = 0$. Hence, $f_1, \ldots, f_n$ is linearly independent.

And given that $\dim V^* = n$, the set $\beta^* = \{f_1, \ldots, f_n\}$ is a basis for V^*. □

Theorem 1.2.25. Let $\beta = \{v_1, \ldots, v_n\}$ be a basis for a vector space V. Then there is a unique dual basis $\beta^* = \{f_1, \ldots, f_n\}$ for V^* such that $f_i(v_j) = \delta_{ij}$.

For each linear functional f on V we have

$$f = \sum_{i=1}^{n} f(v_i) f_i$$

and for each vector $v \in V$ we have

$$v = \sum_{i=1}^{n} f_i(v) v_i$$

Proof. The last proof established that there is a unique basis which is 'dual' to β. Let f be a linear functional on V. Then f is a linear combination of the f_i, so the scalars $c_j = f(v_j)$. Now, if

$$v = \sum_{i=1}^{n} x_i v_i$$

is a vector in V, then

$$f_j(v) = \sum_{i=1}^{n} x_i f_j(v_i) = \sum_{i=1}^{n} x_i \delta_{ij} = x_j$$

so v has a unique expression as a linear combination of v_i given by

$$v = \sum_{i=1}^{n} f_i(v) v_i$$

□

Note that f_i are coordinate functions for β, given that f_i assigns to each vector $v \in V$ the ith coordinate of v relative to the ordered basis β.

How are linear functionals and subspaces related? If f is a non-zero linear functional, then the rank of f is one. And if V is finite-dimensional, then by the Rank–Nullity theorem, the null space N_f has dimension

$$\dim N_f = dim V - 1$$

In a vector space of dimension n, a subspace of of dimension $n-1$ is called a **hyperspace**, which is sometimes called **hyperplanes** or **subspaces of codimension one**. The hyperspace is always the null space of a linear functional.

In the infinite-dimensional case, let W be a proper subspace of V. If there isn't a subspace U such that $W \subsetneq U \subsetneq V$, then W is a hyperplane.

Definition 1.2.15 (Annihilator)**.** Let V be a vector space over $\mathbb{F}$ and S a subset of V. Then the **annihilator** of S is the set S^0 of linear functionals f on V such that $f(v) = 0$ for every $v \in S$.

$$S^0 = \{f \in V^* : f(v) = 0, \ \forall v \in V\}$$

S^0 is a subspace of V^*. If $S = \{0\}$, then $S^0 = V^*$. If $S = V$, then S^0 is the zero subspace of V^*.

The next example shows an important procedure in the following proofs.

Example 1.2.6. Let $\{e_1, e_2, e_3, e_4, e_5\}$ be the standard basis of $\mathbb{R}^5$ and $\{f_1, f_2, f_3, f_4, f_5\}$ be the dual basis of $\mathbb{R}^5$. Suppose

$$W = \text{span}(e_1, e_2) = \{(x_1, x_2, 0, 0, 0) \in \mathbb{R}^5 : x_1, x_2 \in \mathbb{R}\}$$

We show that $W^0 = \text{span}(f_3, f_4, f_5)$.

Recall that f_j is the linear functional that selects the jth coordinate, i.e. $f_j(x_1, x_2, x_3, x_4, x_5) = x_j$.

First suppose $f \in \text{span}(f_3, f_4, f_5)$. Then there exist $c_3, c_4, c_5 \in \mathbb{R}$ such that $f = c_3 f_3 + c_4 f_4 + c_5 f_5$. If $(x_1, x_2, 0, 0, 0) \in W$, then

$$f(x_1, x_2, 0, 0, 0) = (c_3 f_3 + c_4 f_4 + c_5 f_5)(x_1, x_2, 0, 0, 0) = 0$$

Hence $f \in W^0$. I.e., $\text{span}(f_3, f_4, f_5) \subset W^0$.

Now suppose $f \in W^0$. Since the dual basis is a basis of $(\mathbb{R}^5)^*$, there exist $c_1, \ldots, c_5 \in \mathbb{R}$ such that $f = c_1 f_1 + c_2 f_2 + c_3 f_3 + c_4 f_4 + c_5 f_5$. Because $e_1 \in W$ and $f \in W^0$, we have

$$0 = f(e_1) = (c_1 f_1 + c_2 f_2 + c_3 f_3 + c_4 f_4 + c_5 f_5)(e_1) = c_1$$

Similarly, $e_2 \in W$ and thus $c_2 = 0$. Since $e_3, e_4, e_5 \notin W$, $f = c_3 f_3 + c_4 f_4 + c_5 f_5$. Thus $f \in \text{span}(f_3, f_4, f_5)$, i.e., $W^0 \subset \text{span}(f_3, f_4, f_5)$.

The next theorem states that each d-dimensional subspace of an n-dimensional space is the intersection of the null spaces of $(n-d)$ linear functionals.

Theorem 1.2.26. Let V be a finite-dimensional vector space and let W be a subspace of V. Then

$$\dim W + \dim W^0 = \dim V$$

Proof. Let $\{v_1, \ldots, v_k\}$ be a basis for W and choose vectors $\{v_{k+1}, \ldots, v_n \in V$ to extend to a basis $\{v_1, \ldots, v_n\}$ of V. And let $\{f_1, \ldots, f_n\}$ be the basis for V^* which is dual to this basis for V. We show that $\{f_{k+1}, \ldots, f_n\}$ is a basis for W^0.

For $i \geq k+1$, since $f_i(v_j) = \delta_{ij}$ and $\delta_{ij} = 0$ if $i \geq k+1$ and $j \leq k$, we know that f_i belongs to W^0. Hence, for $i \geq k+1$, $f_i(v) = 0$ whenever v is a linear combination of $v_1, \ldots, v_k$.

Given that the functionals $f_{k+1}, \ldots, f_n$ are linearly independent, all we need to show is that they span W^*. Suppose $f \in V^*$. Now

$$f = \sum_{i=1}^{n} f(v_i) f_i$$

implies that if $f \in W^0$, we have $f(v_i) = 0$ for $i \leq k$ and

$$f = \sum_{i=k+1}^{n} f(v_i) f_i$$

Therefore, W^0 has dimension $n-k$, as desired. □

The next corollary shows that if we select some select ordered basis for the space, each k-dimensional subspace can be described by specifying $(n-k)$ homogeneous linear conditions on the coordinates relative to that basis.

Corollary 1.2.27. If W is a k-dimensional subspace of an n-dimensional vector space V, then W is the intersection of $(n-k)$ hyperspaces in V.

Corollary 1.2.28. If W_1 and W_2 are subspaces of a finite-dimensional vector space, then $W_1 = W_2$ iff. $W_1^0 = W_2^0$.

This theory provides a 'dual' point of view on the system of equations, showing how annihilators are related to systems of homogeneous linear equations.

Example 1.2.7. Let W be the subspace of $\mathbb{R}^5$ spanned by the vectors $v_1 = (2,-2,3,4,-1)$, $v_2 = (-1,1,2,5,2)$, $v_3 = (0,0,-1,-2,3)$, and $v_4 = (1,-1,2,3,0)$.

To find the annihilator W^0 of W, we first form a matrix A with row vectors v_1, v_2, v_3, v_4 and find the row-reduced echelon matrix R which is row-equivalent to A.

$$A = \begin{bmatrix} 2 & -2 & 3 & 4 & -1 \\ -1 & 1 & 2 & 5 & 2 \\ 0 & 0 & -1 & -2 & 3 \\ 1 & -1 & 2 & 3 & 0 \end{bmatrix} \longrightarrow R = \begin{bmatrix} 1 & -1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Now, if f is a linear functional on $\mathbb{R}^5$,

$$f(x_1, \ldots, x_5) = \sum_{j=1}^{5} c_j x_j$$

and f is in W^0 iff. $f(v_i) = 0$, for $i = 1,2,3,4$.

This is equivalent to $Ac = 0$, where $c = (c_1, c_2, c_3, c_4, c_5)^t$. Which is, in turn, equivalent to

$Rc = 0$. Or simply

$$\begin{aligned} c_1 - c_2 - c_4 &= 0 \\ c_3 + 2c_4 &= 0 \\ c_5 &= 0 \end{aligned}$$

By setting $c_2 = a$ and $c_4 = b$, we have $c_1 = a + b$, $c_3 = -2b$, $c_5 = 0$. So W^0 consists of all linear functionals of the form

$$f(x_1, x_2, x_3, x_4, x_5) = (a + b)x_1 + ax_2 - 2bx_3 + bx_4$$

The dimension of W^0 is two and a basis $\{f_1, f_2\}$ for it can be found by setting $a = 1, b = 0$, and then $a = 0, b = 1$:

$$\begin{aligned} f_1(x_1, \ldots, x_5) &= x_1 + x_2 \\ f_2(x_1, \ldots, x_5) &= x_1 - 2x_3 + x_4 \end{aligned}$$

And the general form of $f \in W^0$ is $f = af_1 + bf_2$.

1.2.7 Double dual

Bibliography

[HK71] Kenneth Hoffman and Ray Kunze. *Linear Algebra*. 1971.

[YIK89] Manin Yu I and Alexei I Kostrikin. *Linear Algebra and Geometry*. CRC Press, 1989.