Linear Algebra II

Philip Thomas K.

February 18, 2022

Contents

Lecture No. 1

Vector Spaces

1.1 Vector Space Axioms

We define V as a non-empty set of vectors and F as a field of scalars like $\mathbb{R}$, $\mathbb{C}$. Vectors can be added:

$$u + v \qquad u,\ v \in \mathsf{V}.$$

Vectors can be multiplied by scalars:

$$\lambda u \qquad u \in \mathsf{V},\ \lambda \in F.$$

Note. Vector addition or scalar multiplication in general will **not** yield another vector in V.

A vector space follows the following axioms:

1. Closure under vector addition:

$$u + v \in \mathsf{V}, \quad \forall u, v \in \mathsf{V}.$$

2. Closure under scalar multiplication:

$$\lambda u \in \mathsf{V}, \quad \forall u \in \mathsf{V}, \lambda \in F.$$

3. Vector addition is commutative:

$$u + v = v + u, \quad \forall u, v \in \mathsf{V}.$$

4. Vector addition is associative:

$$(u + v) + w = u + (v + w), \quad \forall u, v, w \in \mathsf{V}.$$

5. There is a **unique** zero vector:

$$\exists z \in \mathsf{V} \text{ s.t } v + z = z + v = v, \quad \forall v \in \mathsf{V}.$$

6. There are **unique** negative vectors:
$$\forall v \in \mathsf{V},\ \exists w \in \mathsf{V} \text{ s.t. } v + w = 0.$$

7. Multiplying with the scalar 1 has no effect:
$$1v = v,\ \forall v \in \mathsf{V}.$$

8. Scalar multiplication is associative:
$$(\lambda\mu)v = \lambda(\mu v),\ \forall \lambda, \mu \in F, \forall v \in \mathsf{V}.$$

9. Vector addition is distributive:
$$\lambda(u + v) = \lambda u + \lambda v,\ \forall \lambda \in F, \forall u, v \in \mathsf{V}.$$

10. Addition of scalars is distributive:
$$(\lambda + \mu)v = \lambda v + \mu v,\ \forall \lambda, \mu \in F, \forall v \in \mathsf{V}.$$

To prove that V is a vector space, we must show that all ten axioms are satisfied. To prove that V is not a vector space, we must show that at least one condition is not satisfied.

1.2 Valid Vector Spaces

Example 1.2.1. The set of all n-tuples with entries from a field F is denoted by F^n:
$$\{(a_1, a_2, \ldots, a_n) : a_i \in F\}.$$
if $u = (a_1, a_2, \ldots, a_n) \in F^n$, $v = (b_1, b_2, \ldots, b_n) \in F^n$, and $\lambda \in F$ then:
$$u + v = (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n); \quad \lambda u = (\lambda a_1, \lambda a_2, \ldots, \lambda a_n).$$

Example 1.2.2. $M_{m\times n}(F)$, the set of all $m \times n$ matrices with entries from F:
$$\left\{ \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} : a_{ij} \in F \right\}.$$
The rows of the preceding matrix are regarded as vectors in F^n and the columns are regarded as vectors in F^m. If $A, B \in M_{m\times n}(F)$ and $\lambda \in F$, then:
$$(A + B)_{ij} = a_{ij} + b_{ij}; \quad (\lambda A)_{ij} = \lambda a_{ij},\ (1 \leqslant i \leqslant m,\ 1 \leqslant j \leqslant n).$$

Example 1.2.3. $\mathcal{F}(S, \mathbb{R})$, the set of all functions from S to $\mathbb{R}$, where S is any non-empty set:
$$\{f : S \mapsto \mathbb{R}\}.$$
If $f, g \in \mathcal{F}(S, \mathbb{R})$ and $\lambda \in \mathbb{R}$, then:
$$(f + g)(s) = f(s) + g(s); \quad (\lambda f)(s) = \lambda[f(s)],\ \forall s \in S.$$
These operations are common in calculus.

Example 1.2.4. $P(\mathbb{R})$, the set of all **polynomials** with coefficients from $\mathbb{R}$:

$$\{a_n x^n + \cdots + a_1 x + a_0 : a_i \in \mathbb{R}\}.$$

If $f(x) = a_n x^n + \cdots + a_1 x + a_0$, $g(x) = b_m x^m + \cdots + b_1 x + b_0$, $m \leqslant n$, and $\lambda \in \mathbb{R}$, then:

$$f(x) + g(x) = (a_n + b_n)x^n + \cdots + (a_1 + b_1)x + (a_0 + b_0)$$
$$\lambda f(x) = \lambda a_n x^n + \lambda a_{n-1} x^{n-1} + \cdots + \lambda a_1 x + \lambda a_0$$

1.3 Invalid Vector Spaces

Example 1.3.1. Let $\mathsf{V} = \{(a_1, a_2) : a_1, a_2 \in \mathbb{R}\}$, $F = \mathbb{R}$. For (a_1, a_2), $(b_1, b_2) \in \mathsf{V}$ and $\lambda \in \mathbb{R}$, define the following:

$$(a_1, a_2) + (b_1, b_2) = (a_1 + b_1, a_2 - b_2)$$
$$\lambda(a_1, a_2) = (\lambda a_1, \lambda a_2)$$

Condtions are not satisfied, vector addition is not commutative:

$$(a_1, a_2) + (b_1, b_2) = (a_1 + b_1, a_2 - b_2)$$
$$(b_1, b_2) + (a_1, a_2) = (a_1 + b_1, b_2 - a_2)$$
$$(a_1, a_2) + (b_1, b_2) \neq (b_1, b_2) + (a_1, a_2), \ \forall a_2, b_2 \in \mathbb{R}.$$

Example 1.3.2. Let $\mathsf{V} = \{(a_1, a_2) : a_1, a_2 \in \mathbb{R}\}$, $F = \mathbb{R}$. For (a_1, a_2), $(b_1, b_2) \in \mathsf{V}$ and $\lambda \in \mathbb{R}$, define the following:

$$(a_1, a_2) + (b_1, b_2) = (a_1 + b_1, 0)$$
$$\lambda(a_1, a_2) = (\lambda a_1, 0)$$

Conditions are not satisfied, checking if there is a **unique** zero vector $z = (b_1, b_2)$ such that:

$$u + z = u, \ \forall u = (a_1, a_2) \in \mathsf{V}.$$

Since:

$$\text{LHS} = u + z = (a_1, a_2) + (b_1, b_2) = (a_1 + b_1, 0).$$

the given equality becomes $(a_1 + b_1, 0) = (a_1, a_2)$, $\forall (a_1, a_2) \in \mathsf{V}$, which is equivalent to:

$$z = (b_1, b_2), \ \forall (b_1, b_2) \in \mathsf{V} \text{ s.t. } b_1 = 0.$$

which is not unique.

1.4 Elementary Consequences

The following are elementary consequences of the definition of a vector space:

Cancelation Law for vector addition: $$\forall u, v, w \in \mathsf{V}, \quad u + w = v + w \implies u = v.$$

Properties of Scalar Multiplication in any Vector Space V:

$$0_f u = 0_v, \ \forall u \in \mathsf{V}, \ 0_f \in F, \ 0_v \in \mathsf{V}$$
$$(-\lambda)u = -(\lambda u) = \lambda(-u), \ \forall \lambda \in F, \ \forall u \in \mathsf{V}$$
$$\lambda 0 = 0, \ \forall \lambda \in F, \ 0 \in \mathsf{V}$$

Lecture No. 2

Subspaces

2.1 Basic definition

Abstractly, a subspace is a subset that possesses the same structure as its superset.

Definition 1. Subspace of a vector space. If V is a vector space, and W is a subset of V, then W is a subspace of V, if it is a vector space with respect to addition and scalar multiplication that is defined on V.

In any vector space V, V is also considered a subspace, and $\{0\}$ is also a subspace, called the **zero subspace**.

It is not necessary to verify all 10 conditions of a vector space to prove that a subset of a vector space is a subspace.

Theorem 1. Three condition test for subspaces. Let V be an arbitrary vector space, and let W denote a subset of V. Then W is a subspace iff the following three conditions hold for operations defined in V:

1. $0 \in \mathsf{W}$ (Non-Empty).
2. $u + v \in \mathsf{W}$, $\forall u, v \in \mathsf{W}$ (Closed under vector addition).
3. $\lambda u \in \mathsf{W}$, $\forall \lambda \in F$, $u \in \mathsf{W}$ (Closed under scalar multiplication).

2.2 Examples

Example 2.2.1. Verify if $P_n(\mathbb{R})$, the set of all polynomials in $P(\mathbb{R})$ having degree less than or equal to n $(n \geqslant 0, n \in \mathbb{N})$ is a subspace of $P(\mathbb{R})$.

Axiom 1. $P_n(\mathbb{R})$ is non-empty:

The zero vector in $P_n(\mathbb{R})$ is the zero polynomial, a constant polynomial whose coefficients are all equal to 0. The degree of this polynomial is undefined, but by convention it is set to -1 or $-\infty$. The zero polynomial belongs to $P_n(\mathbb{R})$.

Axiom 2. $P_n(\mathbb{R})$ is closed under vector addition:
Consider two arbitrary polynomials from $P_n(\mathbb{R})$:

$$\begin{aligned} A(x) &= a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \\ B(x) &= b_n x^n + b_{n-1}x^{n-1} + \cdots + b_1 x + b_0 \end{aligned}$$

Their sum is a polynomial of the following form:

$$A(x) + B(x) = (a_n + b_n)x^n + (a_{n-1} + b_{n-1})x^{n-1} + \cdots + (a_1 + b_1)x + (a_0 + b_0).$$

obviously belongs to $P_n(\mathbb{R})$

Axiom 3. $P_n(\mathbb{R})$ is closed under scalar multiplication:

$$\begin{aligned} \lambda A(x) &= \lambda(a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0) \\ &= \lambda a_n x^n + \lambda a_{n-1}x^{n-1} + \cdots + \lambda a_1 x + \lambda a_0 \end{aligned}$$

Which also belongs to $P_n(\mathbb{R})$, making it a subspace of $P(\mathbb{R})$.

Example 2.2.2. Verify if $\mathbf{C}(\mathbb{R})$, the set of all **continuous** real-valued functions defined on $\mathbb{R}$ is a subspace of $\mathcal{F}(\mathbb{R}, \mathbb{R})$.

Axiom 1. $\mathbf{C}(\mathbb{R})$ is non-empty:
The zero vector in $\mathbf{C}(\mathbb{R})$ is the zero function:

$$f_0(x) = 0, \ \forall x \in \mathbb{R}.$$

We write $f_0 \equiv 0$. This function is obviously continuous on $\mathbb{R}$, and therefore $f_0 \in \mathbf{C}(\mathbb{R})$.

Axiom 2. $\mathbf{C}(\mathbb{R})$ is closed under vector addition:
Consider any two functions $f(x), g(x) \in \mathbf{C}(\mathbb{R})$. Their sum is a function of the following form:

$$h(x) = f(x) + g(x).$$

and it is a real-valued function which is continuous on $\mathbb{R}$, $\therefore h(x) \in \mathbf{C}(\mathbb{R})$.

Axiom 3. $\mathbf{C}(\mathbb{R})$ is closed under scalar multiplication:

$$\lambda f(x), \forall \lambda \in \mathbb{R}, f(x) \in \mathbf{C}(\mathbb{R}).$$

Which is obviously in $\mathbf{C}(\mathbb{R})$, making it a subspace of $\mathcal{F}(\mathbb{R}, \mathbb{R})$.

2.3 Forming New Subspaces

The following theorems show how to form a new subspace from other subspaces.

Theorem 2. Any **intersection** of subspaces of a vector space V is a subspace of V.

Proof. Intersections of subspaces. Note that the intersection of subspaces is *not* necessarily finite. Let $\mathcal{S}$ be a collection of subspaces of V, and let W denote the intersection of the subspaces in $\mathcal{S}$.

Axiom 1. W is non-empty:
Since every subspace $\in \mathcal{S}$ contains the zero vector, $0 \in \mathsf{W}$.

Axiom 2. W is closed under vector addition:
Let $u, v \in \mathsf{W}$. Then u, v are contained in each subspace in $\mathcal{S}$. Since each subspace in $\mathcal{S}$ is closed under vector addition, it follows that $u + v$ is contained in each subspace in $\mathcal{S}$.

$$\therefore u, v \in \mathsf{W} \implies u + v \in \mathsf{W}.$$

Axiom 3. W is closed under scalar multiplication:
Let $u \in \mathsf{W}$. Then u is in every subspace in $\mathcal{S}$. Since each subspace in $\mathcal{S}$ is closed under scalar multiplication, it follows that λu is contained in each subspace in $\mathcal{S}$.

$$\therefore u \in \mathsf{W} \implies \lambda u \in \mathsf{W}.$$

So the intersection of subspaces of a vector space is also a subspace. □

Theorem 3. The union of two subspaces of a vector space V is **not necessarily** a subspace of V. The union contains the zero vector, and is closed under scalar multiplication, but in general the *union* of subspaces need not be closed under vector addition.

Proof. Consider a vector space $\mathsf{V} = \mathbb{R}^2$, the Cartesian plane, and two subspaces:

$$\mathsf{W}_1 = \{(x, y) : y = 0\}$$
$$\mathsf{W}_2 = \{(x, y) : x = 0\}$$

This is the x-axis and y-axis, whose union $\mathsf{W}_1 \cup \mathsf{W}_2$ is not a subspace of $\mathbb{R}^2$. □

Lecture No. 3

Linear Combinations & Independence

3.1 Linear Combinations & Span

Definition 2. Linear Combination. Consider a vector space V over a field F and a non-empty subset $S \subset \mathsf{V}$. Then a vector $v \in \mathsf{V}$ is a linear combination of vectors in S if:

$$\exists u_1, u_2, \ldots, u_n \in S, \quad \lambda_1, \lambda_2, \ldots, \lambda_n \in F.$$

such that:

$$v = \lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_n u_n.$$

We can then define span(S) to be the set containing all linear combinations of the vectors in S.

Theorem 4. Smallest subspace. The set span(S) is the smallest subspace of V that contains S. This means that span(S) is a subspace of V that contains S, and that $\forall$ W subspaces of V:

$$S \subseteq \mathsf{W} \implies \text{span}(S) \subseteq \mathsf{W}.$$

Proof. This result is clearly true if $S = \varnothing$, because span$(\varnothing)$=$\{0\}$. Now considering the case where $S \neq \varnothing$, we must first prove that span(S) is a subspace of V containing S. We can easily verify the sufficient conditions since all the linear combinations of the elements in S will be closed under vector addition and scalar multiplication.

Moreover, $S \subset$ span(S), because any vector $u \in S$ can always be written as a linear combination of the vectors in S, namely:

$$u = 1u, \ u \in S.$$

Next we must prove that $\forall$ W subspaces of V:

$$S \subseteq \mathsf{W} \implies \text{span}(S) \subseteq \mathsf{W}.$$

Let v be an arbitrary vector in span(S). This means that:

$$v = \lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_n u_n.$$

where $u_1, u_2, \ldots, u_n \in S$ and $\lambda_1, \lambda_2, \ldots, \lambda_n \in F$. Since $u_1, u_2, \ldots, u_n \in S$, these vectors are in W. Taking into account the fact that W is a subspace of V, we see that:

$$\begin{aligned} \lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_n u_n &\in \mathsf{W} \\ \therefore v &\in \mathsf{W} \\ \therefore \operatorname{span}(S) &\subseteq \mathsf{W} \end{aligned}$$

□

As a consequence of the theorem we can say that a subset of a vector space is a subspace if and only if the set equals its span.

3.2 Linear Dependence & Independence

Definition 3. Linear Dependence & Independence. Consider a vector space V over a field F, and a set $S \subset \mathsf{V}$. The set S is considered to be **linearly dependent** if:

$$\exists u_1, u_2, \ldots, u_n \in S,\ \ \lambda_1, \lambda_2, \ldots, \lambda_n \in F.$$

Where $u_1 \neq u_2 \neq \cdots \neq u_n$ and $\forall n \in \mathbb{Z}, \lambda_n \neq 0$ such that:

$$\lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_n u_n = 0.$$

Conversely we say that the set S is **linearly independent** if the only solution to the equation:

$$\lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_n u_n = 0.$$

is where $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$.

Example 3.2.1. Determine if the set $\{p_0(x), p_1(x), \ldots, p_n(x)\}$, where:

$$\begin{aligned} p_0(x) &= 1 + x + x^2 + \cdots + x^n \\ p_1(x) &= x + x^2 + \cdots + x^n \\ &\ \vdots \\ p_n(x) &= x^n \end{aligned}$$

is linearly independent in $P_n(\mathbb{R})$. Consider the equation:

$$\lambda_0 p_0(x) + \lambda_1 p_1(x) + \cdots + \lambda_n p_n(x) = 0.$$

which is equivalent to the following system of equations:

$$\begin{aligned}
\lambda_0 + \lambda_1 + \cdots + \lambda_{n-1} + \lambda_n &= 0 && \text{(for } x^n\text{)} \\
\lambda_0 + \lambda_1 + \cdots + \lambda_{n-1} &= 0 && \text{(for } x^{n-1}\text{)} \\
&\vdots && \vdots \\
\lambda_0 + \lambda_1 &= 0 && \text{(for x)} \\
\lambda_0 &= 0 && \text{(for the constant term)}
\end{aligned}$$

By back-substitution we get:

$$\lambda_0 = \lambda_1 = \lambda_2 = \cdots = \lambda_{n-1} = \lambda_n = 0.$$

Which shows that the set is linearly independent in $P_n(\mathbb{R})$.

The following is an important result of the definition of linear dependence and independence.

Theorem 5. Consider a vector space V over a field F, and the following relation between sets:

$$S_1 \subset S_2 \subset \mathsf{V}.$$

Then if S_1 is linearly dependent, S_2 is also linearly dependent. And if S_2 is linearly independent, S_1 is also linearly independent.

From this theorem, we can have the following statement on linear dependence.

Theorem 6. Suppose the S is any linearly dependent set such that:

$$|S| \geqslant 2.$$

Then $\exists v \in S$ which can be written as a linear combination of the other vectors in S, and therefore the subset obtained by removing v from S has the same span as S.

Proof. Since S is linearly dependent, $\exists u_1, u_2, \ldots, u_n \in S$ and $\exists \lambda_1, \lambda_2, \ldots, \lambda_n \in F$ where $\forall n \in \mathbb{Z},\ \lambda_n \neq 0$ such that:

$$\lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_n u_n = 0.$$

This gives the following:

$$u_1 = -\frac{\lambda_2}{\lambda_1} u_2 - \cdots - \frac{\lambda_n}{\lambda_1} u_n.$$

which means that u_1 can be expressed as a linear combination of vectors $u_2, \ldots, u_n$, and therefore there will be some $v \in S$ as in the theorem stated. $\square$

Lecture No. 4

Bases & Dimension

4.1 Bases

Definition 4. Basis. Consider a vector space V over a field F. A set $\beta \subset \mathsf{V}$ is a basis for V, if β is both linearly independent and it generates V:

$$\text{span}(\beta) = \mathsf{V}.$$

Example 4.1.1. Because span($\varnothing$)=$\{0\}$ and $\varnothing$ is linearly independent, $\varnothing$ is a basis for the zero vector space.
Example 4.1.2. In F^n the set $\{e_1, e_2, \ldots, e_n\}$, where $e_1 = (1, 0, 0, \ldots, 0), e_2 = (0, 1, 0, \ldots, 0), \ldots, e_n = (0, 0, \ldots, 1)$, is the standard basis for F^n.

Theorem 7. Consider a vector space V over a field F, and the basis for V represented as $\beta = \{u_1, u_2, \ldots, u_n\} \subset \mathsf{V}$ then:

$$\forall v \in \mathsf{V}\ \exists! \lambda_1, \lambda_2, \ldots, \lambda_n \in F \text{ such that } v = \lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_n u_n.$$

Where $\exists!$ means "there uniquely exists".

Proof. Suppose $\beta = \{u_1, u_2, \ldots, u_n\}$ is a basis for V. Now let $v \in \mathsf{V}$, since β generates V, we have $v \in \text{span}(\beta)$. Then there exist $\lambda_1, \lambda_2, \lambda_n \in F$ such that:

$$v = \lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_n u_n.$$

Now assume there is another set of scalars $\mu_1, \mu_2, \ldots, \mu_n$ such that:

$$v = \mu_1 u_1 + \mu_2 u_2 + \cdots + \mu_n u_n.$$

Taking the difference of these two linear combinations we get:

$$(\lambda_1 - \mu_1)u_1 + (\lambda_2 - \mu_2)u_2 + \cdots + (\lambda_n - \mu_n)u_n = 0.$$

However, β is linearly independent, and so the coefficients in the equation above are all zero:

$$\therefore \lambda_1 = \mu_1, \lambda_2 = \mu_2, \ldots, \lambda_n = \mu_n.$$

Thus v is uniquely expressed as a linear combination of the vectors in β. $\square$

Theorem 8. Consider a vector space V over a field F, and a finite set S that generates V. Then some subset of S is a basis for V, and V has a finite basis.

Based on this theorem, we have a method for reducing a finite spanning set to a finite basis. Let $\{u_1, u_2, \dots, u_n\}$ be a finite spanning set of vectors in V with $u_1 \neq 0$.

Step 1: Choose $u_1 \neq 0$ and keep it in the 'expected' basis.

Step 2: Determine whether u_2 is a linear combination of the remaining vectors to its left.

- If it is, then cross off u_2.
- If it isn't then keep u_2.

Step k: Determine whether u_k is a linear combination of the remaining vectors to its left.

- If it is, then cross off u_k.
- If it isn't then keep u_k.

Step n: Determine whether u_n is a linear combination of the remaining vectors to its left.

- If it is, then cross off u_n, and the remaining vectors to its let form a basis for V.
- If it isn't then keep u_n, and the remaining vectors to its left together with u_n form a basis for V.

4.2 Dimension

Theorem 9. Replacement Theorem. Consider a vector space V over a field F, a set $G \subset \mathsf{V}$ that generates V and contains n vectors, and a linearly independent set $L \subset \mathsf{V}$ that contains m vectors. Then:

$$m \leqslant n.$$

and:

$$\exists G' \subset G,\ |G'| = n - m \text{ such that } L \cup G' \text{ generates } \mathsf{V}.$$

As a result of this theorem, we can say that if V, a vector space, has a finite basis, then every basis for V contains the same number of vectors.

Proof. Suppose β is a basis for V containing n vectors. Now, let γ be any other basis for V. Now assume that γ contains more that n vectors (at least $n+1$). Then we can select a set $S \subset \gamma$ containing $n+1$ vectors which are all linearly independent.

Now we have:

$$S \text{ is linearly independent, } |S| = n+1,$$
$$\beta \text{ generates } \mathsf{V},\ |\beta| = n.$$

And by the first conclusion of the replacement theorem:

$$|S| \leqslant |\beta| \equiv n+1 \leqslant n.$$

Which is a contradiction. Therefore, γ is finite, and the number m of vectors in γ satisfies $m \leqslant n$. Interchanging the roles of β and γ, and arguing as above, we obtain $n \leqslant m$, therefore $m = n$. □

This fact lends itself to the definition of the dimension of a vector space.

Definition 5. Dimension, finite-dimensional, & infinite-dimensional. When the basis β of a vector space V over a field F has a cardinality of n:

$$|\beta| = n.$$

We say that the **dimension** of V is n:

$$\dim(\mathsf{V}) = n.$$

If n is finite, we say that V is **finite-dimensional**. If n is not finite, we say that V **infinite-dimensional**.

Example 4.2.1. Standard dimensions of vector spaces:

1. $\dim(\{0\}) = 0$.
2. $\dim(F^n) = n$.
3. $\dim(M_{m\times n}(F)) = mn$.
4. $\dim(P_n(F)) = n+1$.

Example 4.2.2. The dimension of a vector space depends on its field of scalars. If $\mathsf{V} = \mathbb{C}$ and $F = \mathbb{C}$ then:

$$\dim(\mathbb{C}(\text{over } \mathbb{C})) = 1.$$

And an appropriate basis would be $\beta = \{1\}$.
If $\mathsf{V} = \mathbb{C}$ and $F = \mathbb{R}$ then:

$$\dim(\mathbb{C}(\text{over } \mathbb{R})) = 2.$$

And an appropriate basis would be $\beta = \{1, i\}$.

If you have a linearly independent subset $\mathsf{W} \subset \mathsf{V}$ which has a cardinality lower than the dimension of V, you can extend it to create a basis for V by making a set $\mathsf{W}' = \mathsf{W} + \beta$ where β is the standard basis for V, and apply the reduction algorithm in section 4.1.

Lecture No. 5

Linear Transformation Notation, Null Space, & Range

5.1 Basic Notation

If V and W are vector spaces, the mapping T from V to W is a function that assigns to each vector $v \in \mathsf{V}$ a unique vector $w \in \mathsf{W}$. In this case we say that T maps V into W, and write:

$$\mathsf{T} : \mathsf{V} \mapsto \mathsf{W}.$$

For each $v \in \mathsf{V}$ the vector $w = \mathsf{T}(v) \in \mathsf{W}$ is the image of v under T.

Definition 6. Linear Transformation. Consider the vector spaces V and W over a field F. Then $\mathsf{T} : \mathsf{V} \mapsto \mathsf{W}$ is a **linear transformation** if $\forall u, v \in \mathsf{V}$, $\forall \lambda \in F$ the following conditions hold:

$$\begin{aligned} \mathsf{T}(u+v) &= \mathsf{T}(u) + \mathsf{T}(v) \\ \mathsf{T}(\lambda u) &= \lambda \mathsf{T}(u), \end{aligned}$$

or:

$$\mathsf{T}(\lambda u + \mu v) = \lambda \mathsf{T}(u) + \mu \mathsf{T}(v), \forall u, v \in \mathsf{V},\ \forall \lambda, \mu \in F.$$

Example 5.1.1. Determine if the transformation $\mathsf{T} : \mathbb{R}^2 \mapsto \mathbb{R}^2$ defined by:

$$\mathsf{T}(a_1, a_2) = (a_1 + a_2, -a_1).$$

is linear.
Solution. Writing the vectors in column form:

$$\mathsf{T} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} a_1 + a_2 \\ -a_1 \end{bmatrix}.$$

Consider the arbitrary vectors:

$$u = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix},\ v = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \in \mathbb{R}^2,$$

and scalars $\lambda, \mu \in \mathbb{R}$. We must verify the condition for linearity. For the left hand side we have:

$$\begin{aligned} \mathsf{T}(\lambda u + \mu v) &= \mathsf{T}\left(\begin{bmatrix} \lambda a_1 \\ \lambda a_2 \end{bmatrix} + \begin{bmatrix} \mu b_1 \\ \mu b_2 \end{bmatrix}\right) \\ &= \mathsf{T}\begin{bmatrix} \lambda a_1 + \mu b_1 \\ \lambda a_2 + \mu b_2 \end{bmatrix} \\ &= \begin{bmatrix} (\lambda a_1 + \mu b_1) + (\lambda a_2 + \mu b_2) \\ -(\lambda a_1 + \mu b_1) \end{bmatrix} \end{aligned}$$

For the right hand side we have:

$$\begin{aligned} \lambda \mathsf{T}(u) + \mu \mathsf{T}(v) &= \lambda \begin{bmatrix} a_1 + a_2 \\ -a_1 \end{bmatrix} + \mu \begin{bmatrix} b_1 + b_2 \\ -b_1 \end{bmatrix} \\ &= \begin{bmatrix} (\lambda a_1 + \mu b_1) + (\lambda a_2 + \mu b_2) \\ -(\lambda a_1 + \mu b_1) \end{bmatrix} \end{aligned}$$

The left hand side and the right hand side are equal and therefore T is a linear transformation.

Example 5.1.2. Determine if the transformation $\mathsf{T} : \mathsf{M}_{2\times 2}(\mathbb{R}) \mapsto \mathbb{R}$ defined by:

$$\mathsf{T}(A) = \det(A).$$

or otherwise written:

$$\mathsf{T}\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} = a_1 a_4 - a_2 a_3.$$

is linear.

Solution. In this case it is sufficient to show that either of the two conditions for linearity are not satisfied. Looking at the condition regarding scalar multiplication:

$$\begin{aligned} \mathsf{T}(\lambda A) &= \lambda a_1 \lambda a_4 - \lambda a_2 \lambda a_3 \\ &= \lambda^2 \det(A) \end{aligned}$$

and on the other hand we have:

$$\lambda \mathsf{T}(A) = \lambda \det(A).$$

Both sides of the equation are not equivalent, and therefore the transformation is not linear.

Theorem 10. Let $\mathsf{T} : \mathsf{V} \mapsto \mathsf{W}$ be a linear transformation. The following properties hold as a consequence:

1. $\mathsf{T}(0) = 0$, or $\mathsf{T}(0_v) = 0_w$
2. $\mathsf{T}(\lambda u + v) = \lambda \mathsf{T}(u) + \mathsf{T}(v)$, $\forall u, v \in \mathsf{V}$, $\forall \lambda \in F$
3. $\mathsf{T}(u - v) = \mathsf{T}(u) - \mathsf{T}(v)$, $\forall u, v \in \mathsf{V}$
4. $\forall u_1, u_2, \ldots, u_n \in \mathsf{V}$, $\forall \lambda_1, \lambda_2, \ldots, \lambda_n \in F$:

$$\mathsf{T}\left(\sum_{i=1}^{n} \lambda_i u_i\right) = \sum_{i=1}^{n} \lambda_i \mathsf{T}(u_i).$$

Note. The second property combines the two requirements for the linearity of T into one statement, and is generally used to prove that a transformation is linear.
We can use linear transformations to describe some familiar concepts. Differentiation can be described as $\mathsf{T} : P_n(\mathbb{R}) \mapsto P_{n-1}(\mathbb{R})$:

$$\mathsf{T}(f(x)) = f'(x), \ \forall f \in P_n(\mathbb{R}).$$

Proof. Let $g(x), h(x) \in P_n(\mathbb{R})$ and $\lambda \in \mathbb{R}$. Consider:

$$\begin{aligned} \mathsf{T}(\lambda g(x) + h(x)) &= (\lambda g(x) + h(x))' \\ &= \lambda g'(x) + h'(x) \\ &= \lambda \mathsf{T}(g(x)) + \mathsf{T}(h(x)) \end{aligned}$$

And therefore T is a linear transformation. $\square$

5.2 Null Space & Range

There are two basic and important transformations that appear frequently called the **identity** and **zero** transformations.

Definition 7. Identity & Zero transformation. The **identity** transformation $I_{\mathsf{V}} : \mathsf{V} \mapsto \mathsf{V}$ is defined by:

$$I_{\mathsf{V}}(u) = u, \ \forall u \in \mathsf{V}.$$

The **zero** transformation $\mathsf{T}_0 : \mathsf{V} \mapsto \mathsf{W}$ is defined by:

$$\mathsf{T}_0(u) = 0, \ \forall u \in \mathsf{V}.$$

They are both linear transformations.

There are also two important sets associated with linear transformations. They are called the **null space** and **range**.

Definition 8. Null Space (Kernel) & Range (Image). Consider a linear transformation $\mathsf{T} : \mathsf{V} \mapsto \mathsf{W}$. $N(\mathsf{T})$ the **null space** (or **kernel**) of T is defined as:

$$N(\mathsf{T}) = \{v \in \mathsf{V} : \mathsf{T}(v) = 0\}.$$

And $R(\mathsf{T})$ the **range** (or **image**) of T is defined as:

$$R(\mathsf{T}) = \{\mathsf{T}(v) : v \in \mathsf{V}\}.$$

Thus the null space is the set of all vectors in V that are mapped to the zero vector, while the range is the set of all images in W of the mapping.

Example 5.2.1. For the identity transformation $I_{\mathsf{V}} : \mathsf{V} \mapsto \mathsf{V}$:

$$\begin{aligned} N(I_{\mathsf{V}}) &= \{0\} \\ R(I_{\mathsf{V}}) &= \mathsf{V} \end{aligned}$$

For the zero transformation $\mathsf{T}_0 : \mathsf{V} \mapsto \mathsf{W}$:

$$N(\mathsf{T}_0) = \mathsf{V}$$
$$R(\mathsf{T}_0) = \{0\}$$

Example 5.2.2. Find the null space and range of the transformation $\mathsf{T} : \mathbb{R}^3 \mapsto \mathbb{R}^2$ defined by:

$$\mathsf{T}(a_1, a_2, a_3) = (a_1 - a_2, 2a_3).$$

Solution. For $N(\mathsf{T})$ we have:

$$\mathsf{T}\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} a_1 - a_2 \\ 2a_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Which implies the following:

$$a_1 = a_2$$
$$a_3 = 0$$

and so we can describe the null space as such:

$$N(\mathsf{T}) = \{(a, a, 0) : a \in \mathbb{R}\}.$$

As $R(\mathsf{T}) \subseteq \mathbb{R}^2$, it suffices to prove that $\mathbb{R}^2 \subseteq R(\mathsf{T})$. So we must prove that for any $v = (b_1, b_2) \in \mathbb{R}^2$ there exists $w = (a_1, a_2, a_3) \in \mathbb{R}^3$ such that $\mathsf{T}(w) = v$:

$$\mathsf{T}\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} a_1 - a_2 \\ 2a_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}.$$

Which implies that:

$$a_1 - a_2 = b_1$$
$$2a_3 = b_2$$

Which is a set of linear equations that have infinitely many solutions, one of which is $a_1 = b_1,\ a_2 = 0,\ a_3 = \frac{b_2}{2}$ and hence $R(\mathsf{T}) = \mathbb{R}^2$

We see that in the examples above, the null space and range of each of the linear transformations are a subspace. This leads to the next theorem.

Theorem 11. Let T be a linear transformation that maps the vector space V to the vector space W:

$$\mathsf{T} : \mathsf{V} \mapsto \mathsf{W}.$$

Then, the null space $N(\mathsf{T})$ and $R(\mathsf{T})$ are subspaces of V and W respectively.

Proof. Null Space of a Linear Transformation is a Subspace. For a set of vectors to be a subspace, we must check the three axioms are satisfied:
Axiom 1. $N(\mathsf{T})$ is non-empty: Since $\mathsf{T}(0_\mathsf{V}) = 0_\mathsf{W}, 0_\mathsf{V} \in N(\mathsf{T})$.

Axiom 2. $N(\mathsf{T})$ is closed under vector addition: We see that $\forall u, v \in N(\mathsf{T})$ we have:

$$\begin{aligned} \mathsf{T}(u+v) &= \mathsf{T}(u) + \mathsf{T}(v) \\ &= 0_{\mathsf{W}} + 0_{\mathsf{W}} \\ &= 0_{\mathsf{W}} \\ \therefore (u+v) &\in N(\mathsf{T}) \quad \forall u, v \in N(\mathsf{T}). \end{aligned}$$

Axiom 3. $N(\mathsf{T})$ is closed under scalar multiplication: We see that $\forall u \in N(\mathsf{T})$ and $\forall \lambda \in F$ we have:

$$\begin{aligned} \mathsf{T}(\lambda u) &= \lambda \mathsf{T}(u) \\ &= \lambda \cdot 0_{\mathsf{W}} \\ &= 0_{\mathsf{W}} \\ \therefore \lambda u &\in N(\mathsf{T}) \quad \forall u \in N(\mathsf{T}),\ \lambda \in F \end{aligned}$$

The above method can be applied to prove that the range is also a subspace. □

For a given linear transformation, we can find $N(\mathsf{T})$ by simply solving the equation:

$$\mathsf{T}(v) = 0.$$

Where the set of all solutions is precisely $N(\mathsf{T})$, which in practice reduces to solving a linear system of equations. To find $R(\mathsf{T})$, we can find the basis for the range we can apply the transformation to the basis.

Theorem 12. Let T be a linear transformation:

$$\mathsf{T} : \mathsf{V} \mapsto \mathsf{W}.$$

And let β be a basis for V:

$$\beta = \{v_1, v_2, \dots, v_n\}.$$

Then:

$$R(\mathsf{T}) = \text{span}(\mathsf{T}(\beta)).$$

Proof. Basis for Range is the Linear Transformation of Basis Vectors. First we prove that $R(\mathsf{T}) \subseteq \text{span}(\mathsf{T}(\beta))$. Take an arbitrary vector $w \in R(\mathsf{T})$. The fact that $w \in R(\mathsf{T})$ implies that there is $v \in \mathsf{V}$ with $w = \mathsf{T}(v)$. Since β is a basis for V, we can represent v in terms of vectors in β:

$$v = \sum_{i=1}^{n} \lambda_i v_i, \qquad \lambda_i \in F \ (1 \leqslant i \leqslant n).$$

Hence, by the linearity of T:

$$w = \mathsf{T}(v) = \mathsf{T}\left(\sum_{i=1}^{n} \lambda_i v_i\right) = \sum_{i=1}^{n} \lambda_i \mathsf{T}(v_i) \in \text{span}(\mathsf{T}(\beta)).$$

So $R(\mathsf{T}) \subseteq \text{span}(\mathsf{T}(\beta))$.

Next we prove $\text{span}(\mathsf{T}(\beta)) \subseteq R(\mathsf{T})$. Take an arbitrary vector $w \in \text{span}(\mathsf{T}(\beta))$. Then there exists scalars $\lambda_1, \lambda_2, \ldots, \lambda_n \in F$ such that:

$$\begin{aligned} w &= \lambda_1 \mathsf{T}(v_1) + \lambda_2 \mathsf{T}(v_2) + \cdots + \lambda_n \mathsf{T}(v_n) \\ &= \mathsf{T}(\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n) \end{aligned}$$

Therefore, w is the image under T of a linear combination of vectors in V. This shows that $w \in R(\mathsf{T})$, and $\therefore \text{span}(\mathsf{T}(\beta)) \subseteq R(\mathsf{T})$. $\square$

Example 5.2.3. The linear transformation $\mathsf{T} : P_2(\mathbb{R}) \mapsto \mathsf{M}_{2\times 2}(\mathbb{R})$ is defined by:

$$\mathsf{T}(f(x)) = \begin{bmatrix} f(1) - f(2) & 0 \\ 0 & f(0) \end{bmatrix}.$$

Find a basis for $R(\mathsf{T})$ and $\dim(R(\mathsf{T}))$.
Solution. Since $\beta = \{1, x, x^2\}$ is a basis for $P_2(\mathbb{R})$, we have:

$$\begin{aligned} R(\mathsf{T}) &= \text{span}(\mathsf{T}(\beta)) \\ &= \text{span}\left(\{\mathsf{T}(1), \mathsf{T}(x), \mathsf{T}\left(x^2\right)\}\right) \\ &= \text{span}\left(\left\{\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} -3 & 0 \\ 0 & 0 \end{bmatrix}\right\}\right) \end{aligned}$$

Notice that $\begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} -3 & 0 \\ 0 & 0 \end{bmatrix}$ are linearly dependent. Then:

$$\begin{aligned} R(\mathsf{T}) &= \text{span}\left(\left\{\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} -3 & 0 \\ 0 & 0 \end{bmatrix}\right\}\right) \\ &= \text{span}\left(\left\{\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}\right\}\right) \end{aligned}$$

Since the two vectors in the resultant set are linearly independent, we can conclude that it is a basis for $R(\mathsf{T})$, and so $\dim(R(\mathsf{T})) = 2$.

Lecture No. 6

Rank & Nullity

6.1 Rank-Nullity Theorem

The null space and range are important and have special names attached to them.

Definition 9. Nullity & Rank. Let T be a linear transformation $\mathsf{T} : \mathsf{V} \mapsto \mathsf{W}$, where $\dim(N(\mathsf{T})) < \infty$ and $\dim(R(\mathsf{T})) < \infty$. Then,

$$\dim(N(\mathsf{T}))$$

is known as the nullity of T. And then,

$$\dim(R(\mathsf{T}))$$

is known as the rank of T.

The balance between rank and nullity is reflected in the following theorem.

Theorem 13. Rank-Nullity Theorem. Let T be a linear transformation $\mathsf{T} : \mathsf{V} \mapsto \mathsf{W}$ and let $\dim(\mathsf{V}) < \infty$. Then:

$$\text{nullity}(\mathsf{T}) + \text{rank}(\mathsf{T}) = \dim(\mathsf{V}).$$

Otherwise written:

$$\dim(N(\mathsf{T})) + \dim(R(\mathsf{T})) = \dim(\mathsf{V}).$$

Proof. Rank-Nullity Theorem. Suppose that $\dim(\mathsf{V}) = n$ and $\dim(N(\mathsf{T})) = k$. Then we consider three cases.
Case 1: $0 < k < n$, that is $k \in \{1, 2, \ldots, n-1\}$.

Let $\{v_1, v_2, \ldots, v_k\}$ be a basis for $N(\mathsf{T})$. Therefore this basis for $N(\mathsf{T})$ is a linearly independent subset of V. Then by the **Replacement Theorem**, we may extend $\{v_1, v_2, \ldots, v_k\}$ to a basis for V:

$$\beta = \{v_1, v_2, \ldots, v_k, v_{k+1}, \ldots, v_n\}.$$

Now consider the following set S:

$$S = \{\mathsf{T}(v_{k+1}), \mathsf{T}(v_{k+2}), \dots, \mathsf{T}(v_n)\}.$$

And consider the fact that:

$$R(\mathsf{T}) = \operatorname{span}(\{\mathsf{T}(v_1), \mathsf{T}(v_2), \dots, \mathsf{T}(v_n)\}).$$

Since $\mathsf{T}(v_i) = 0$ for $1 \leqslant i \leqslant k$ because $v_i \in N(\mathsf{T})$ we have:

$$\begin{aligned} R(\mathsf{T}) &= \operatorname{span}(\{\mathsf{T}(v_{k+1}), \mathsf{T}(v_{k+2}), \dots, \mathsf{T}(v_n)\}) \\ &= \operatorname{span}(S). \end{aligned}$$

Now to prove that S is linearly independent, we consider the equation:

$$a_{k+1}\mathsf{T}(v_{k+1}) + a_{k+2}\mathsf{T}(v_{k+2}) + \cdots + a_n\mathsf{T}(v_n) = 0, \quad a_{k+1}, a_{k+2}, \dots, a_n \in F.$$

And since T is linear, we can rewrite the equation as:

$$\mathsf{T}(a_{k+1}v_{k+1} + a_{k+2}v_{k+2} + \cdots + a_n v_n) = 0.$$

Therefore:

$$a_{k+1}v_{k+1} + a_{k+2}v_{k+2} + \cdots + a_n v_n \in N(\mathsf{T}).$$

Since $\{v_1, v_2, \dots, v_k\}$ is a basis for $N(\mathsf{T})$, there exist $c_1, c_2, \dots, c_k \in F$ such that:

$$a_{k+1}v_{k+1} + a_{k+2}v_{k+2} + \cdots + a_n v_n = c_1v_1 + c_2v_2 + \cdots + c_kv_k.$$

Otherwise rewritten:

$$-c_1v_1 - c_2v_2 - \cdots - c_kv_k + a_{k+1}v_{k+1} + a_{k+2}v_{k+2} + \cdots + a_n v_n = 0.$$

And since β is a basis for V and is linearly independent, the coefficients of the last equation must all be 0. Thus S is a linearly independent set and is a basis for $R(\mathsf{T})$, and therefore:

$$\operatorname{rank}(\mathsf{T}) = \dim(R(\mathsf{T})) = n - k.$$

<u>Case 2</u>: $k = n$.

In this case, the image of every vector in V is the zero vector in W, so that $R(\mathsf{T}) = \{0\}$, and therefore $\dim(R(\mathsf{T})) = \dim(\{0\}) = 0$. The statement of the theorem is true for this case.

<u>Case 3</u>: $k = 0$. In this case $N(\mathsf{T}) = \{0\}$, so the nullity is 0. If $\{v_1, v_2, \dots, v_n\}$ is a basis for V, then by the theorem above we have:

$$R(\mathsf{T}) = \operatorname{span}(\{\mathsf{T}(v_1), \mathsf{T}(v_2), \dots, \mathsf{T}(v_n)\}).$$

And a similar argument to the one above shows that:

$$\{\mathsf{T}(v_1), \mathsf{T}(v_2), \dots, \mathsf{T}(v_n)\}$$

is linearly independent. Thus $\dim(R(\mathsf{T})) = n = \dim(\mathsf{V})$, and the result also holds in this case. $\square$

Example 6.1.1. $\mathsf{T} : P_4(\mathbb{R}) \mapsto P_2(\mathbb{R})$, linear, defined by:

$$\mathsf{T}(p(x)) = p^{(2)}(x).$$

Find a basis for $N(\mathsf{T})$ as well as for $R(\mathsf{T})$.
Solution. For $N(\mathsf{T})$, by definition:

$$p(x) = a + bx + cx^2 + dx^3 + ex^4 \in N(\mathsf{T}) \iff \mathsf{T}(p(x)) = 0.$$

Or equivalently:

$$p^{(2)}(x) = 2c + 6dx + 12ex^2 = 0, \ \forall x \in \mathbb{R} \iff c = d = e = 0.$$

Thus:

$$p \in N(\mathsf{T}) \iff p(x) = a + bx, \ a, b \in \mathbb{R}.$$

This shows that $N(\mathsf{T})$ consists of all polynomials of degree at most 1. So $N(\mathsf{T}) = P_1(\mathbb{R})$. Hence, $\{1, x\}$ is a basis for $N(\mathsf{T})$, and $\dim(N(\mathsf{T})) = 2$.

For $R(\mathsf{T})$: since $\dim(P_4(\mathbb{R})) = 5$, by the rank-nullity theorem, we have:

$$2 + \dim(R(\mathsf{T})) = 5 \implies \dim(R(\mathsf{T})) = 3.$$

Now we consider the standard basis $\beta = \left\{1, x, x^2, x^3, x^4\right\}$ for $P_4(\mathbb{R})$ to get:

$$\begin{aligned} R(\mathsf{T}) &= \operatorname{span}\left(\left\{\mathsf{T}(1), \mathsf{T}(x), \mathsf{T}(x^2), \mathsf{T}(x^3), \mathsf{T}(x^4)\right\}\right) \\ &= \operatorname{span}\left(\left\{0, 0, 2, 6x, 12x^2\right\}\right) \\ &= \operatorname{span}\left(\left\{2, 6x, 12x^2\right\}\right) \end{aligned}$$

Three vectors generate $R(\mathsf{T})$, and hence since $\dim(R(\mathsf{T})) = 3$, these vectors form a basis for $R(\mathsf{T})$. Observe that $R(\mathsf{T})$ is just $P_2(\mathbb{R})$.

Lecture No. 7

Properties of Linear Transformations

7.1 One-to-one and onto

For further study of properties of linear transformations, the concept of **one-to-one** and **onto** mappings is needed.

Definition 10. One-to-one, onto, and bijective. A mapping T is **one-to-one** (or **injective**), if:

$$x \neq y \implies \mathsf{T}(x) \neq \mathsf{T}(y) \iff \mathsf{T}(x) = \mathsf{T}(y) \implies x = y.$$

That is, distinct elements of V must have distinct images in W. T is **onto** (or **surjective**), if:

$$\mathsf{T}(\mathsf{V}) = \mathsf{W}.$$

That is, the range of T is W. T is **bijective** if it is both injective and surjective.

Example 7.1.1. Let $\mathsf{T} : \mathbb{R}^2 \mapsto \mathbb{R}^2$ be defined by:

$$\mathsf{T}(v) = Av, \quad A = \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}.$$

Show that T is one-to-one and onto.
Solution. To show that T is one-to-one, let:

$$u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \qquad v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}.$$

Then:

$$\mathsf{T}(u) = \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \qquad \mathsf{T}(v) = \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

Now if $\mathsf{T}(u) = \mathsf{T}(v)$, then:

$$\begin{bmatrix} u_1 + u_2 \\ -u_1 \end{bmatrix} = \begin{bmatrix} v_1 + v_2 \\ -v_1 \end{bmatrix}.$$

The last matrix equation gives $u_1 = v_1$, $u_2 = v_2$, that is, $u = v$. Thus T is one-to-one. To show that T is onto, take an arbitrary $w = \begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R}^2$. We must show that there is a vector $v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \in \mathbb{R}^2$ such that:

$$\mathsf{T}(v) = \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix} \iff \begin{bmatrix} v_1 + v_2 \\ -v_1 \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}.$$

The last matrix equation gives $v_1 = -b$, $v_2 = a + b$. Thus T is onto.
For a linear transformation, both of these concepts are connected to the rank and nullity of the transformation. The following result gives a useful way to detrmine whether a linear transformation is one-to-one.

Theorem 14. Consider the linear transformation:

$$\mathsf{T} : \mathsf{V} \mapsto \mathsf{W}.$$

Then the followng bijective statement holds true:

$$\mathsf{T} \text{ one-to-one} \iff N(\mathsf{T}) = \{0\}.$$

That is, the null space only contains the trivial solution.

Proof. T one-to-one $\iff N(\mathsf{T}) = \{0\}$. Suppose that T is one-to-one and $v \in N(\mathsf{T})$. Now, we must show that $v = 0$. Indeed:

$$\begin{rcases} v \in N(\mathsf{T}) \implies \mathsf{T}(v) = 0 \\ \mathsf{T} \text{ linear} \implies \mathsf{T}(0) = 0 \end{rcases} \implies v = 0.$$

Since T is one-to-one. Thus $N(\mathsf{T}) = \{0\}$. Now, suppose that $\mathsf{T}(u) = \mathsf{T}(v)$. Then $0 = \mathsf{T}(u) - \mathsf{T}(v)$. By linearity of T, we have $\mathsf{T}(u) - \mathsf{T}(v) = \mathsf{T}(u - v)$. Then we get $\mathsf{T}(u - v) = 0$, which means that $u - v \in N(\mathsf{T}) = \{0\}$. So $u - v = 0$ or $u = v$. Thus T is one-to-one. ☐

Example 7.1.2. Consider the linear transformation $\mathsf{T} : \mathbb{R}^2 \mapsto \mathbb{R}^2$ defined by:

$$\mathsf{T} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x - 3y \\ 5x + 2y \end{bmatrix}.$$

Use the theorem above to prove that T is one-to-one.

Solution. The vector $\begin{bmatrix} x \\ y \end{bmatrix}$ is in $N(\mathsf{T})$ if and only if:

$$\begin{aligned} 2x - 3y &= 0 \\ 5x + 2y &= 0 \end{aligned}$$

This linear system has the unique solution $x = y = 0$. Thus $N(\mathsf{T}) = \{0\}$ and hence by the theorem above, T is one-to-one.

In general, a linear transformation may be one-to-one without being onto and may be onto without being one-to-one. Surprisingly, these properties are equivalent in an important special case.

Theorem 15. Consider the linear transformation $\mathsf{T} : \mathsf{V} \mapsto \mathsf{W}$ where:

$$\dim(\mathsf{V}) = \dim(\mathsf{W}) < \infty.$$

Then the following are equivalent:

1. T is **one-to-one**
2. T is **onto**
3. $\operatorname{rank}(\mathsf{T}) = \dim(\mathsf{V})$, that is $\dim(R(\mathsf{T})) = \dim(\mathsf{V})$

Proof. Linear transformations that preserve dimension are bijective. We use the **dimension theorem**:

$$\dim(N(\mathsf{T})) + \dim(R(\mathsf{T})) = \dim(\mathsf{V}).$$

By the theorem above:

$$\begin{aligned}
\mathsf{T} \text{ is one-to-one} &\iff N(\mathsf{T}) = \{0\} \\
&\iff \dim(N(\mathsf{T})) = 0 \\
&\iff \dim(R(\mathsf{T})) = \dim(\mathsf{V}) \iff \dim(R(\mathsf{T})) = \dim(\mathsf{W}) \\
&\iff \dim(R(\mathsf{T})) = \dim(\mathsf{W})
\end{aligned}$$

Since $R(\mathsf{T}) \subseteq \mathsf{W}$, the last equality is equivalent to $R(\mathsf{T}) = \mathsf{W}$, which means that T is onto. $\square$

Example 7.1.3. Consider the linear transformation $\mathsf{T} : P_2(\mathbb{R}) \mapsto P_3(\mathbb{R})$ defined by:

$$\mathsf{T}(f(x)) = 2f'(x) + \int_0^x 3f(t)\,\mathrm{d}t.$$

Solution. Since $\{1, x, x^2\}$ is a basis for $P_2(\mathbb{R})$, we have:

$$\begin{aligned}
R(\mathsf{T}) &= \operatorname{span}\left(\{\mathsf{T}(1), \mathsf{T}(x), \mathsf{T}\left(x^2\right)\}\right) \\
&= \operatorname{span}\left(\left\{3x, 2 + \frac{3}{2}x^2, 4x + x^3\right\}\right)
\end{aligned}$$

We can confirm that $\left\{3x, 2 + \frac{3}{2}x^2, 4x + x^3\right\}$ is linearly independent, then:

$$\operatorname{rank}(\mathsf{T}) = \dim(R(\mathsf{T})) = 3.$$

Because $\dim(P_3(\mathbb{R})) = 4$, T is not onto. From the dimension theorem, $\dim(N(\mathsf{T})) + 3 = 3$, so $\dim(N(\mathsf{T})) = 0$, and therefore, $N(\mathsf{T}) = \{0\}$. This means that T is one-to-one.

7.2 Uniqueness of linear transformations

One of the most important properties of a linear transformation is that it is completely determined by its action on a basis.

Theorem 16. Unique Linear Transformation. Suppose that we have vector spaces V & W over the field F. And suppose that V has a basis:

$$\{v_1, v_2, \ldots, v_n\}.$$

Then for every set $w_1, w_2, \ldots, w_n \in \mathsf{W}$, there exists uniquely a linear transformation $\mathsf{T} : \mathsf{V} \mapsto \mathsf{W}$ depending on the set such that:

$$\mathsf{T}(v_i) = w_i, \ \forall i = 1, 2, \ldots, n.$$

Proof. Existence of unique linear transformation. For this proof, we have to prove two parts, the existence and the uniqueness of T. To prove existence, we let $v \in \mathsf{V}$. Then v can be represented uniquely in the form:

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

where $\lambda_1, \lambda_2, \ldots, \lambda_n$ are scalars. Then, let us define $\mathsf{T} : \mathsf{V} \mapsto \mathsf{W}$ by the rule:

$$\mathsf{T}(v) = \sum_{i=1}^{n} \lambda_i w_i.$$

We can prove that T is linear. Indeed, for $u, v \in \mathsf{V}$ and $\lambda \in F$ there are scalars $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$ such that:

$$u = \sum_{i=1}^{n} a_i v_i \quad \text{and} \quad v = \sum_{i=1}^{n} b_i v_i.$$

Then $\lambda u + v = \sum_{i=1}^{n} (\lambda a_i + b_i) v_i$, and so

$$\begin{aligned}
\mathsf{T}(\lambda u + v) &= \mathsf{T}\left(\sum_{i=1}^{n} (\lambda a_i + b_i) v_i\right) \\
&= \sum_{i=1}^{n} (\lambda a_i + b_i) w_i \\
&= \lambda \sum_{i=1}^{n} a_i w_i + \sum_{i=1}^{n} b_i w_i \\
&= \lambda \mathsf{T}(u) + \mathsf{T}(v).
\end{aligned}$$

Thus it is clear that:

$$\mathsf{T}(v_i) = w_i \ \forall i = 1, 2, \ldots, n.$$

To prove the transformation's uniqueness, we first suppose that there exists another linear transformation $\mathsf{U} : \mathsf{V} \mapsto \mathsf{W}$ such that $\mathsf{U}(v_i) = w_i$ for $i = 1, 2, \ldots, n$. Then for $v \in \mathsf{V}$ with $v = \sum_{i=1}^{n} \lambda_i v_i$, we have:

$$\mathsf{U}(v) = \sum_{i=1}^{n} \lambda_i \mathsf{U}(v_i) = \sum_{i=1}^{n} \lambda_i w_i = \mathsf{T}(v).$$

Thus $\mathsf{U}(v) = \mathsf{T}(v)$, $\forall v \in \mathsf{V}$, that is $\mathsf{U} = \mathsf{T}$. $\square$

Example 7.2.1. Let $\mathsf{T} : \mathbb{R}^2 \mapsto \mathbb{R}^2$ be a linear transformation defined by:

$$\mathsf{T}(a_1, a_2) = (2a_2 - a_1, 3a_1).$$

And let $\mathsf{U} : \mathbb{R}^2 \mapsto \mathbb{R}^2$ be a linear transformation.

If we know that $\mathsf{U}(1,2) = (3,3)$ and $\mathsf{U}(1,1) = (1,3)$, then $\mathsf{U} = \mathsf{T}$. This follows from the fact that $\{(1,2),(1,1)\}$ is a basis for $\mathbb{R}^2$.

Lecture No. 8

Matrix Representation of Linear Transformations

Suppose that we have a basis $\beta = \{v_1, v_2, \ldots, v_n\}$ for a vector space V over F. Then for every vector $v \in \mathsf{V}$ there are unique scalars $c_1, c_2, \ldots, c_n \in F$, such that:

$$v = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n.$$

Here, we are trying to associate the list of scalars $\{c_1, c_2, \ldots, c_n\}$ with the basis vectors in β. Note that changing the order of the basis β will change the order of the scalars.

Example 8.0.1. The set of two vectors:

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

form a basis for $\mathbb{R}^2$, but for the two bases:

$$\beta = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\} \quad \text{and} \quad \beta' \left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\},$$

we have:

$$\begin{aligned} \begin{bmatrix} 1 \\ 2 \end{bmatrix} &= 1 \cdot \beta_1 + 2 \cdot \beta_2 \\ &= 2 \cdot \beta_1' + 1 \cdot \beta_2' \end{aligned}$$

Thus the list of scalars associated with the vector is $\{1, 2\}$ relative to β, and $\{2, 1\}$ relative to β'. To deal with this ambiguity, we introduce the concept of an **ordered basis** for a vector space.

> **Definition 11.** Ordered Basis. Suppose we have a vector space V such that $\dim(\mathsf{V}) < \infty$. Then an **ordered basis** for V is a basis for V in a specific order, or equivalently, a finite sequence of vectors in V which is linearly independent and spans V.

For the vector space F^n, we call:

$$\left\{ e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \ldots, e_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \right\}$$

the standard ordered basis for F^n. For the vector space $P_n(F)$, we call:

$$\{1, x, x^2, \ldots, x^n\}$$

the standard ordered basis for $P_n(F)$.

With the concept of ordered bases, we can identify abstract vectors in an n-dimensional vector space with n-tuples.

Definition 12. Co-ordinate Vector. Let $\beta = \{u_1, u_2, \ldots, u_n\}$ be an ordered basis for V, and let $v \in \mathsf{V}$ and $c_1, c_2, \ldots, c_n$ be the unique scalars such that:

$$v = c_1 u_1 + c_2 u_2 + \cdots + c_n u_n.$$

We define the **co-ordinate vector** of v relative to β, denoted by $[v]_\beta$ to be:

$$[v]_\beta = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}.$$

Example 8.0.2. Suppose we have $\mathsf{V} = P_2(\mathbb{R})$ and $f(x) = 1 - 2x + 5x^2$. Then, if we choose $\beta = \{1, x, x^2\}$ to be the standard ordered basis for V, then we have:

$$[f]_\beta = \begin{bmatrix} 1 \\ -2 \\ 5 \end{bmatrix}.$$

However, if we choose $\beta' = \{1, x+1, (x+1)^2\}$ to be an alternate ordered basis for V, then find $[f]_{\beta'}$.
Solution. We must find $a, b, c \in \mathbb{R}$ such that:

$$a(1) + b(x+1) + c(x+1)^2 = 1 - 2x + 5x^2,$$

or equivalently,

$$(a+b+c) + (b+2c)x + cx^2 = 1 - 2x + 5x^2.$$

This gives us the following system of linear equations:

$$\begin{aligned} a + b + c &= 1 \\ b + 2c &= -2 \\ c &= 5 \end{aligned}$$

Whose solution is $a = 8, b = -12, c = 5$. Therefore:

$$[f]_{\beta'} = \begin{bmatrix} 8 \\ -12 \\ 5 \end{bmatrix}.$$

Suppose that we have two finite dimensional vector spaces. V with an ordered basis $\beta = \{v_1, v_2, \ldots, v_n\}$, and W with an ordered basis $\gamma = \{w_1, w_2, \ldots, w_m\}$. Suppose then that we have a transformation $\mathsf{T} : \mathsf{V} \mapsto \mathsf{W}$ that is linear. If this is true, then $\forall j \in \{1, 2, \ldots, n\}$, there exist unique scalars $c_{1j}, c_{2j}, \ldots, c_{mj} \in F$ such that:

$$\begin{aligned} \mathsf{T}(v_1) &= c_{11}w_1 + c_{21}w_2 + \cdots + c_{m1}w_m \\ \mathsf{T}(v_2) &= c_{12}w_1 + c_{22}w_2 + \cdots + c_{m2}w_m \\ &\vdots \\ \mathsf{T}(v_n) &= c_{1n}w_1 + c_{2n}w_2 + \cdots + c_{mn}w_m \end{aligned}$$

We can write the equations in a matrix form:

$$\begin{bmatrix} \mathsf{T}(v_1) \\ \mathsf{T}(v_2) \\ \vdots \\ \mathsf{T}(v_n) \end{bmatrix} = \begin{bmatrix} c_{11} & c_{21} & \cdots & c_{m1} \\ c_{12} & c_{22} & \cdots & c_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ c_{1n} & c_{2n} & \cdots & c_{mn} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{bmatrix}.$$

Definition 13. Matrix Representation. The matrix:

$$A = (c_{ij})_{m\times n} = \begin{bmatrix} c_{11} & c_{21} & \cdots & c_{m1} \\ c_{12} & c_{22} & \cdots & c_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ c_{1n} & c_{2n} & \cdots & c_{mn} \end{bmatrix}$$

is called the **matrix representation** of T with respect to the ordered bases β and γ and we write $A = [\mathsf{T}]_\beta^\gamma$. If $\mathsf{V} = \mathsf{W}$ and $\beta = \gamma$, then we write $A = [\mathsf{T}]_\beta$.

Note. The matrix A is in fact the **transpose** of the left matrix in the matrix product above.

Example 8.0.3. Consider the linear transformation $\mathsf{T} : \mathbb{R}^2 \mapsto \mathbb{R}^3$ defined by:

$$\mathsf{T}(a_1, a_2) = (a_1 + 3a_2, 0, 2a_1 - 4a_2).$$

If we choose β and γ to be the standard ordered bases for $\mathbb{R}^2$ and $\mathbb{R}^3$ respectively, then:

$$\begin{aligned} \mathsf{T}(1,0) &= (1,0,2) = 1e_1 + 0e_2 + 2e_3 \\ \mathsf{T}(0,1) &= (3,0,-4) = 3e_1 + 0e_2 - 4e_3 \end{aligned} \implies [\mathsf{T}]_\beta^\gamma = \begin{bmatrix} 1 & 3 \\ 0 & 0 \\ 2 & -4 \end{bmatrix}.$$

Moving on, let V & W be vector spaces over F. And let $\mathsf{T}, \mathsf{U} : \mathsf{V} \mapsto \mathsf{W}$ be arbitrary transformations. Then, we define $\mathsf{T} + \mathsf{U} : \mathsf{V} \mapsto \mathsf{W}$ by:

$$(\mathsf{T} + \mathsf{U})(v) = \mathsf{T}(v) + \mathsf{U}(v), \ \forall v \in \mathsf{V},$$

and $\lambda\mathsf{T} : \mathsf{V} \mapsto \mathsf{W}$ by:

$$(\lambda\mathsf{T})(v) = \lambda\mathsf{T}(v), \ \forall v \in \mathsf{V}.$$

With linear transformations, we can preserve linearity.

Theorem 17. Vector space of all linear transformations. Let V and W be vector spaces over the field F. And let $\mathsf{T}, \mathsf{U} : \mathsf{V} \mapsto \mathsf{W}$ be linear transformations. Then:

$$\forall \lambda \in F, \ \lambda\mathsf{T} + \mathsf{U} \text{ is a linear transformation.}$$

The set of all linear transformations $\mathsf{V} \mapsto \mathsf{W}$ is a vector space over F with addition and scalar multiplication as defined above.

The vector space of all linear transformations $\mathsf{V} \mapsto \mathsf{W}$ over the same field F is denoted by $\mathcal{L}(\mathsf{V}, \mathsf{W})$. If $\mathsf{V} = \mathsf{W}$, then we write $\mathcal{L}(\mathsf{V})$ instead.

Theorem 18. Let V and W be finite-dimensional vector spaces, with ordered bases β and γ. And let $\mathsf{T}, \mathsf{U} : \mathsf{V} \mapsto \mathsf{W}$ with $\mathsf{T}, \mathsf{U} \in \mathcal{L}(\mathsf{V}, \mathsf{W})$. Then the following statements are true:

$$[\mathsf{T} + \mathsf{U}]_\beta^\gamma = [\mathsf{T}]_\beta^\gamma + [\mathsf{U}]_\beta^\gamma$$
$$[\lambda\mathsf{T}]_\beta^\gamma = \lambda[\mathsf{T}]_\beta^\gamma, \ \forall \lambda \in F$$

Example 8.0.4. Let $\mathsf{S}, \mathsf{T} \in \mathcal{L}\left(\mathbb{R}^2\right)$ be defined by:

$$\mathsf{S}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + 2y \\ -y \end{bmatrix} \quad \text{and} \quad \mathsf{T}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -x + y \\ 3x \end{bmatrix}.$$

If β is the standard basis for $\mathbb{R}^2$, find $[\mathsf{S} + \mathsf{T}]_\beta$ and $[3\mathsf{S}]_\beta$ by using the definition and by using the theorem above.

Solution. The matrix representations for S and T are:

$$[\mathsf{S}]_\beta = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix} \quad \text{and} \quad [\mathsf{T}]_\beta = \begin{bmatrix} -1 & 1 \\ 3 & 0 \end{bmatrix},$$

respectively. By definition, we have:

$$\begin{aligned} (\mathsf{S} + \mathsf{T})\begin{bmatrix} x \\ y \end{bmatrix} &= \mathsf{S}\begin{bmatrix} x \\ y \end{bmatrix} + \mathsf{T}\begin{bmatrix} x \\ y \end{bmatrix} \\ &= \begin{bmatrix} x + 2y \\ -y \end{bmatrix} + \begin{bmatrix} -x + y \\ 3x \end{bmatrix} \\ &= \begin{bmatrix} 3y \\ 3x - y \end{bmatrix} \end{aligned}$$

and therefore, its matrix representation is:

$$[(\mathsf{S}+\mathsf{T})]_\beta = \begin{bmatrix} 0 & 3 \\ 3 & -1 \end{bmatrix}.$$

Similarly,

$$\begin{aligned}(3\mathsf{S})\begin{bmatrix} x \\ y \end{bmatrix} &= 3\mathsf{S}\begin{bmatrix} x \\ y \end{bmatrix} \\ &= 3\begin{bmatrix} x+2y \\ -y \end{bmatrix} \\ &= \begin{bmatrix} 3x+6y \\ -3y \end{bmatrix}\end{aligned}$$

and hence:

$$[3\mathsf{S}]_\beta = \begin{bmatrix} 3 & 6 \\ 0 & -3 \end{bmatrix}.$$

By the theorem, we have:

$$[\mathsf{S}+\mathsf{T}]_\beta = [\mathsf{S}]_\beta + [\mathsf{T}]_\beta = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix} + \begin{bmatrix} -1 & 1 \\ 3 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 3 \\ 3 & -1 \end{bmatrix}.$$

and:

$$[3\mathsf{S}]_\beta = 3[\mathsf{S}]_\beta = 3\begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 3 & 6 \\ 0 & -3 \end{bmatrix}.$$

Lecture No. 9

Compositions of Linear Transformations

Definition 14. Composition. Let A, B, and C be sets and $f : A \mapsto B$ and $g : B \mapsto C$ be transformations. The **composition** of g and f, denoted by $g \circ f$, is a function $g \circ f : A \mapsto C$, defined as follows:

$$(g \circ f)(x) = g(f(x)), \ \forall x \in A.$$

The first result states that the composition of linear transformations is linear.

Theorem 19. Let V, W, and Z be vector spaces over the field F. If we have some $\mathsf{T} \in \mathcal{L}(\mathsf{V}, \mathsf{W})$ and some $\mathsf{U} \in \mathcal{L}(\mathsf{W}, \mathsf{Z})$, then the following statement holds true:

$$\mathsf{U} \circ \mathsf{T} \in \mathcal{L}(\mathsf{V}, \mathsf{Z}).$$

Proof. If $u, v \in \mathsf{V}$ and $\lambda \in F$, then:

$$\begin{aligned}
(\mathsf{U} \circ \mathsf{T})(\lambda u + v) &= \mathsf{U}(\mathsf{T}(\lambda u + v)) \\
&= \mathsf{U}(\lambda \mathsf{T}(u) + \mathsf{T}(v)) \\
&= \lambda \mathsf{U}(\mathsf{T}(u)) + \mathsf{U}(\mathsf{T}(v)) \\
&= \lambda (\mathsf{U} \circ \mathsf{T})(u) + (\mathsf{U} \circ \mathsf{T})(v)
\end{aligned}$$

□

Theorem 20. Let V, W, and Z be finite-dimensional vector spaces with ordered bases α, β, and γ, respectively. We have the following implication:

$$\mathsf{T} \in \mathcal{L}(\mathsf{V}, \mathsf{W}), \mathsf{U} \in \mathcal{L}(\mathsf{W}, \mathsf{Z}) \implies [\mathsf{U} \circ \mathsf{T}]_\alpha^\gamma = [\mathsf{U}]_\beta^\gamma [\mathsf{T}]_\alpha^\beta.$$

Corollary 9.0.1. Given V with the ordered basis β and some $\mathsf{T}, \mathsf{U} \in \mathcal{L}(\mathsf{V})$, we have:

$$[\mathsf{U} \circ \mathsf{T}]_\beta = [\mathsf{U}]_\beta [\mathsf{T}]_\beta.$$

Example 9.0.1. Consider the transformations $\mathsf{U} \in \mathcal{L}(P_3(\mathbb{R}), P_2(\mathbb{R}))$ and $\mathsf{T} \in \mathcal{L}(P_2(\mathbb{R}), P_3(\mathbb{R}))$ defined by:

$$\mathsf{U}(f(x)) = f'(x), \qquad \mathsf{T}(f(x)) = \int_0^x f(t)\,\mathrm{d}t,$$

with α and β representing the standard ordered bases of $P_3(\mathbb{R})$ & $P_2(\mathbb{R})$ respectively. Show that $\mathsf{U} \circ \mathsf{T} = I$, the identity transformation on $P_2(\mathbb{R})$.
Solution. First, recall that $\alpha = \{1, x, x^2, x^3\}$ and $\beta = \{1, x, x^2\}$. We have:

$$[\mathsf{U}]_\alpha^\beta = \left[[\mathsf{U}(1)]_\beta, [\mathsf{U}(x)]_\beta, \left[\mathsf{U}\left(x^2\right)\right]_\beta, \left[\mathsf{U}\left(x^3\right)\right]_\beta\right] = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix},$$

and:

$$[\mathsf{T}]_\beta^\alpha = \left[[\mathsf{T}(1)]_\alpha, [\mathsf{T}(x)]_\alpha, \left[\mathsf{T}\left(x^2\right)\right]_\alpha\right] = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix}.$$

Therefore:

$$[\mathsf{U} \circ \mathsf{T}]_\beta = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = [I]_\beta.$$

Next, we show how to evaluate the transformation at any vector point.

Theorem 21. Let V, W be finite dimensional vector spaces with the ordered bases β, and γ respectively. Also let $\mathsf{T} \in \mathcal{L}(\mathsf{V}, \mathsf{W})$, then:

$$[\mathsf{T}(u)]_\gamma = [\mathsf{T}]_\beta^\gamma [u]_\beta, \ \forall u \in V.$$

Proof. Let $u \in \mathsf{V}$, and let us define the following linear transformations:

$$\begin{aligned} &\mathsf{R}(\lambda) = \lambda u & &\mathsf{R} : F \mapsto \mathsf{V} & &\forall \lambda \in F \\ &\mathsf{S}(\lambda) = \lambda \mathsf{T}(u) & &\mathsf{S} : F \mapsto \mathsf{W} & &\forall \lambda \in F \end{aligned}$$

Then $\mathsf{S} = \mathsf{T} \circ \mathsf{R}$, since $\forall \lambda \in F$ we have $\mathsf{R}(\lambda) = \lambda u$, and so:

$$\mathsf{T}(\mathsf{R}(\lambda)) = \mathsf{T}(\lambda u) = \lambda \mathsf{T}(u) = \mathsf{S}(\lambda).$$

Let $\alpha = \{1\}$ be the standard ordered basis for F. Applying theorem 20 to R, S, and T, we have:

$$[\mathsf{T} \circ \mathsf{R}]_\alpha^\gamma = [\mathsf{T}]_\beta^\gamma [\mathsf{R}]_\alpha^\beta.$$

However:

$$\begin{aligned}
[\mathsf{T}(u)]_\gamma &= [1\mathsf{T}(u)] \\
&= [\mathsf{S}(1)]_\gamma && \text{(since we have: } \mathsf{S}(\lambda) = \lambda\mathsf{T}(u)) \\
&= [\mathsf{S}]_\alpha^\gamma && \text{(since } \alpha = \{1\} \text{ is a basis for } F) \\
&= [\mathsf{T} \circ \mathsf{R}]_\alpha^\gamma && \text{(since } \mathsf{S} = \mathsf{T} \circ \mathsf{R}) \\
&= [\mathsf{T}]_\beta^\gamma [\mathsf{R}]_\alpha^\beta && \text{(by theorem 20)} \\
&= [\mathsf{T}]_\beta^\gamma [\mathsf{R}(1)]_\beta && \text{(since } \alpha = \{1\} \text{ is a basis for } F) \\
&= [\mathsf{T}]_\beta^\gamma [u]_\beta && \text{(since } \mathsf{R}(\lambda) = \lambda u).
\end{aligned}$$

□

Definition 15. Left Multiplication Transformation. Let A be an $m \times n$ matrix with entries from a field F. We denote L_A to be the mapping $\mathsf{L}_A : F^n \mapsto F^m$ defined by:

$$\mathsf{L}_A(x) = Ax, \text{ for each column vector } x \in F^n.$$

We call L_A a **left-multiplication transformation**.

Example 9.0.2. Let

$$A = \begin{bmatrix} 1 & 3 & 5 \\ 6 & 4 & 2 \end{bmatrix}.$$

Then $A \in \mathsf{M}_{2\times 3}(\mathbb{R})$ and $\mathsf{L}_A : \mathbb{R}^3 \mapsto \mathbb{R}^2$. If

$$x = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \in \mathbb{R}^3,$$

then

$$\mathsf{L}_A(x) = Ax = \begin{bmatrix} 1 & 3 & 5 \\ 6 & 4 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} -4 \\ 4 \end{bmatrix} \in \mathbb{R}^2.$$

Theorem 22. Let $A \in \mathsf{M}_{m\times n}(F)$. Then $\mathsf{L}_A \in \mathcal{L}(F^n, F^m)$. Furthermore, if $B \in \mathsf{M}_{m\times n}(F)$, and β and γ are the standard ordered bases for F^n and F^m respectively, then the following are equivalent:

(a) $[\mathsf{L}_A]_\beta^\gamma = A$

(b) $\mathsf{L}_A = \mathsf{L}_B \iff A = B$

(c) $\mathsf{L}_{A+B} = \mathsf{L}_A + \mathsf{L}_B$ and $\mathsf{L}_{\lambda A} = \lambda \mathsf{L}_A$, $\forall \lambda \in F$

(d) $\forall \mathsf{T} \in \mathcal{L}(F^n, F^m)$, $\exists! C_{m\times n} \ni \mathsf{T} = \mathsf{L}_C$, where $C = [T]_\beta^\gamma$

(e) $E \in \mathsf{M}_{n\times p}(F) \implies \mathsf{L}_{AE} = \mathsf{L}_A \mathsf{L}_E$

(f) $m = n \implies \mathsf{L}_{I_n} = I_{F^n}$

Proof. The fact that L_A is linear follows from the properties of matrix operations. Let $\beta = \{e_1, e_2, \ldots, e_n\}$.

(a) We already noted that the j-th column of the $m \times n$ matrix $[\mathsf{L}_A]_\beta^\gamma$ is just $[\mathsf{L}_A(e_j)]_\gamma$, the coordinate vector of $\mathsf{L}_A(e_j)$ in terms of γ. However, by definition, $\mathsf{L}_A(e_j) = Ae_j$ (a product of the two matrices), which is precisely the j-th column of the matrix A. So $[\mathsf{L}_A]_\beta^\gamma = A$.

(b) Suppose that $\mathsf{L}_A = \mathsf{L}_B$. Then by (a), we can write $A = [\mathsf{L}_A]_\beta^\gamma$ and $B = [\mathsf{L}_B]_\beta^\gamma$. Hence $A = B$ and the converse implication is trivial.

(c) By definition, for any $v \in F^n$

$$\mathsf{L}_{A+B}(v) = (A+B)v = Av + Bv = \mathsf{L}_A(v) + \mathsf{L}_B(v) = (\mathsf{L}_A + \mathsf{L}_B)(v).$$

(d) Suppose that $\mathsf{T} \in \mathcal{L}(F^n, F^m)$. Take $C = [T]_\beta^\gamma$. Then we have

$$[\mathsf{T}(v)]_\gamma = [\mathsf{T}]_\beta^\gamma [v]_\beta, \ \forall v \in F^n,$$

which, in our situation, is

$$[\mathsf{T}(v)]_\gamma = C[v]_\beta, \ \forall v \in F^n.$$

Note that since β and γ are the standard bases for F^n and F^m respectively, we have

$$[v]_\beta = v, \ \forall v \in F^n \qquad [w]_\gamma = w, \ \forall w \in F^m.$$

Then the relation

$$[\mathsf{T}(v)]_\gamma = C[v]_\beta, \ \forall v \in F^n$$

turns into

$$\mathsf{T}(v) = Cv.$$

And, by definition, $Cv = \mathsf{L}_C(v)$, $\forall v \in F^n$. Thus $\mathsf{T} = \mathsf{L}_C$, and the existence of C is proven. The uniqueness of C follows from (b).

(e) Let $\{e'_1, e'_2, \ldots, e'_p\}$ be the standard basis for F^p. Since matrix multiplication is associative, we have $(AE)e'_j = A(Ee'_j)$ $\forall j \in \{1, 2, \ldots, p\}$. Thus

$$\begin{aligned} \mathsf{L}_{AE}(e'_j) &= (AE)e'_j = A(Ee'_j) = \mathsf{L}_A(Ee'_j) = \mathsf{L}_A(\mathsf{L}_E(e'_j)) \\ &= (\mathsf{L}_A \circ \mathsf{L}_E)(e'_j) \end{aligned}$$

which means that $\mathsf{L}_{AE} = \mathsf{L}_A \circ \mathsf{L}_E$.

(f) Suppose that $m = n$. Recall that I_n is the $n \times n$ **identity matrix**, and I_{F^n} is the **identity linear transformation** from F^n into F^n. Then, $\forall v \in F^n$ we have

$$\mathsf{L}_{I_n}(v) = I_n v = v = I_{F^n}(v),$$

which shows that $\mathsf{L}_{I_n} = I_{F^n}$.

∎

Lecture No. 10

Invertibility & Isomorphisms

Definition 16. Let V and W be vector spaces, and let $\mathsf{T} \in \mathcal{L}(\mathsf{V}, \mathsf{W})$. Then a function $\mathsf{U} : \mathsf{W} \mapsto \mathsf{V}$ is said to be an **inverse** of T if:

$$\mathsf{T} \circ \mathsf{U} = I_{\mathsf{W}} \quad \text{and} \quad \mathsf{U} \circ \mathsf{T} = I_{\mathsf{V}}.$$

If T has an inverse, then T is said to be **invertible**.

We know from `MH1300` that if T is invertible, then its inverse is unique and is denoted by T^{-1}. We also know that it implies that T is one-to-one and onto.

Theorem 23. Let V and W be finite-dimensional vector spaces where $\dim(\mathsf{V}) = \dim(\mathsf{W})$, and let $\mathsf{T} \in \mathcal{L}(\mathsf{V}, \mathsf{W})$. Then:

$$\mathsf{T} \text{ invertible} \iff \dim(R(\mathsf{T})) = \dim(\mathsf{V}).$$

Since invertibility implies one-to-one and onto.

The following result shows that the inverse of a linear transformation preserves linearity.

Theorem 24. Given an invertible $\mathsf{T} \in \mathcal{L}(\mathsf{V}, \mathsf{W})$ we conclude that $\mathsf{T}^{-1} : \mathsf{W} \mapsto \mathsf{V}$ is also linear, that is, $\mathsf{T}^{-1} \in \mathcal{L}(\mathsf{W}, \mathsf{V})$.

Proof. Let $w_1, w_2 \in \mathsf{W}$ and $\lambda \in F$. Since T is onto and one-to-one, there exists unique vectors $v_1, v_2 \in \mathsf{V}$ such that:

$$\mathsf{T}(v_1) = w_1 \quad \text{and} \quad \mathsf{T}(v_2) = w_2.$$

Thus:

$$v_1 = \mathsf{T}^{-1}(w_1) \quad \text{and} \quad v_2 = \mathsf{T}^{-1}(w_2),$$

and we have:

$$\begin{aligned}\mathsf{T}^{-1}(\lambda w_1 + w_2) &= \mathsf{T}^{-1}\left(\lambda \mathsf{T}(v_1) + \mathsf{T}(v_2)\right) \\ &= \mathsf{T}^{-1}\left(\mathsf{T}(\lambda v_1 + v_2)\right) \\ &= \lambda v_1 + v_2 \\ &= \lambda \mathsf{T}^{-1}(w_1) + \mathsf{T}^{-1}(w_2)\end{aligned}$$

□

Lemma 10.0.1. Let $\mathsf{T} \in \mathcal{L}(\mathsf{V}, \mathsf{W})$ be an invertible transformation. Then:

$$\dim(\mathsf{V}) < \infty \iff \dim(\mathsf{W}) < \infty.$$

In this case, $\dim(\mathsf{V}) = \dim(\mathsf{W})$.

Proof. Suppose that V is finite-dimensional. Let $\beta = \{v_1, v_2, \ldots, v_n\}$ be a basis for V. Then $\mathsf{T}(\beta)$ spans $R(\mathsf{T}) = \mathsf{W}$, hence some subset of $\mathsf{T}(\beta)$ is a basis for W, that is, W is finite-dimensional. The reverse implication can be proven by a similar argument using T^{-1}.

Now, suppose that both V and W are finite-dimensional. Because T is one-to-one and onto, we have:

$$\text{nullity}(\mathsf{T}) = \dim(N(\mathsf{T})) = \dim(\{0\}) = 0,$$

$$\text{rank}(\mathsf{T}) = \dim(R(\mathsf{T})) = \dim(\mathsf{W}).$$

So by the dimension theorem,

$$\text{nullity}(\mathsf{T}) + \text{rank}(\mathsf{T}) = \dim(\mathsf{V}) \iff 0 + \dim(\mathsf{W}) = \dim(\mathsf{V}),$$

that is,

$$\dim(\mathsf{V}) = \dim(\mathsf{W}).$$

□

Theorem 25. Let $\mathsf{T} \in \mathcal{L}(\mathsf{V}, \mathsf{W})$, where V and W are finite-dimensional vector spaces with ordered bases β and γ respectively. Then:

$$\mathsf{T} \text{ invertible} \iff [\mathsf{T}]_\beta^\gamma \text{ invertible}.$$

Furthermore:

$$\left[\mathsf{T}^{-1}\right]_\gamma^\beta = \left([\mathsf{T}]_\beta^\gamma\right)^{-1}.$$

Proof. Suppose that T is invertible. This implies that $\dim(\mathsf{V}) = \dim(\mathsf{W})$. Let $n = \dim(\mathsf{V})$. Then $[\mathsf{T}]_\beta^\gamma$ is an $n \times n$ matrix. Next, for $\mathsf{T}^{-1} : \mathsf{W} \mapsto \mathsf{V}$, we have $\mathsf{T} \circ \mathsf{T}^{-1} = I_\mathsf{W}$ and $\mathsf{T}^{-1} \circ \mathsf{T} = I_\mathsf{V}$. Thus:

$$I_n = \begin{cases} [I_\mathsf{V}]_\beta = \left[\mathsf{T}^{-1} \circ \mathsf{T}\right]_\beta = \left[\mathsf{T}^{-1}\right]_\gamma^\beta [\mathsf{T}]_\beta^\gamma \\ [I_\mathsf{W}]_\gamma = \left[\mathsf{T} \circ \mathsf{T}^{-1}\right]_\gamma = [\mathsf{T}]_\beta^\gamma \left[\mathsf{T}^{-1}\right]_\gamma^\beta \end{cases} \implies [\mathsf{T}]_\beta^\gamma \text{ invertible,}$$

and

$$\left([\mathsf{T}]_\beta^\gamma\right)^{-1} = [\mathsf{T}^{-1}]_\gamma^\beta.$$

Now suppose that $A = [\mathsf{T}]_\beta^\gamma$ is invertible. Then $\exists B \in \mathsf{M}_{n\times n}$ such that $AB = BA = I_n$. We apply theorem 16 to a starting vector space W with an ordered basis $\gamma = \{\gamma_1, \gamma_2, \ldots, \gamma_n\}$ and a destination vector space V with the set of vectors:

$$\begin{aligned} v_1 &= b_{11}\beta_1 + b_{21}\beta_2 + \cdots + b_{n1}\beta_n \\ v_2 &= b_{12}\beta_1 + b_{22}\beta_2 + \cdots + b_{n2}\beta_n \\ &\vdots \\ v_n &= b_{1n}\beta_1 + b_{2n}\beta_2 + \cdots + b_{nn}\beta_n \end{aligned}$$

where $\beta = \{\beta_1, \beta_2, \ldots, \beta_n\}$, an ordered basis for V. Then $\exists \mathsf{U} \in \mathcal{L}(\mathsf{W}, \mathsf{V})$, such that:

$$\mathsf{U}(\gamma_k) = v_k = b_{1k}\beta_1 + b_{2k}\beta_2 + \cdots + b_{nk}\beta_n, \quad k = 1, 2, \ldots, n.$$

If follows that $[\mathsf{U}]_\gamma^\beta = B = (b_{ik})$. To show $\mathsf{U} = \mathsf{T}^{-1}$, note that:

$$[\mathsf{U} \circ \mathsf{T}]_\beta = [\mathsf{U}]_\gamma^\beta [\mathsf{T}]_\beta^\gamma = BA = I_n = [I_\mathsf{V}]_\beta.$$

So $\mathsf{U} \circ \mathsf{T} = I_\mathsf{V}$, and similarly, $\mathsf{T} \circ \mathsf{U} = I_\mathsf{W}$. Thus T is invertible. □

Definition 17. Isomorphic. Let V and W be vector spaces. We say that V is **isomorphic** to W if $\exists \mathsf{T} \in \mathcal{L}(\mathsf{V}, \mathsf{W})$ that is **invertible**, and denote:

$$\mathsf{V} \simeq \mathsf{W}.$$

Such a linear transformation is called an **isomorphism** from V onto W.

Example 10.0.1. The vector space $\mathsf{M}_{2\times 2}(\mathbb{R})$ is isomorphic to $\mathbb{R}^4$, because the transformation

$$\mathsf{T}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = (a, b, c, d),$$

is linear, one-to-one, and onto.

Theorem 26. Let V and W be finite-dimensional vector spaces over the same field. Then:

$$\mathsf{V} \simeq \mathsf{W} \iff \dim(\mathsf{V}) = \dim(\mathsf{W}).$$

Proof. Suppose that $\mathsf{V} \simeq \mathsf{W}$ and that $\mathsf{T} : \mathsf{V} \mapsto \mathsf{W}$ is an isomorphism from V onto W. Then we have $\dim(\mathsf{V}) = \dim(\mathsf{W})$. Now suppose that $\dim(\mathsf{V}) = \dim(\mathsf{W})$, and let $\beta = \{v_1, v_2, \ldots, v_n\}$ and $\gamma = \{w_1, w_2, \ldots, w_n\}$ be bases for V and W respectively. By theorem 16 we can say that $\exists \mathsf{T} \in \mathcal{L}(\mathsf{V}, \mathsf{W})$ such that $\mathsf{T}(v_i) = w_i$ for $i = 1, 2, \ldots, n$. Then we have:

$$R(\mathsf{T}) = \operatorname{span}(\mathsf{T}(\beta)) = \operatorname{span}(\gamma) = \mathsf{W}.$$

So T is onto. Then T is also one-to-one, and hence T is an isomorphism. □

The following theorem shows how the collection $\mathcal{L}(\mathsf{V}, \mathsf{W})$ of all linear transformations my be identified with the appropriate vector space of $m \times n$ matrices.

Theorem 27. Let V and W be vector spaces where:

$$\dim(\mathsf{V}) = n \quad \text{and} \quad \dim(\mathsf{W}) = m,$$

and where β and γ are the ordered bases for V and W respectively. Then the transformation $\Phi : \mathcal{L}(\mathsf{V}, \mathsf{W}) \mapsto \mathsf{M}_{m\times n}(F)$ defined by:

$$\Phi(\mathsf{T}) = [\mathsf{T}]_\beta^\gamma, \ \forall \mathsf{T} \in \mathcal{L}(\mathsf{V}, \mathsf{W}),$$

is an **isomorphism**.

Proof. By theorem 18, Φ is linear. All that's left to show is that Φ is one-to-one and onto. This can be achieved by demonstrating that $\forall A \in \mathsf{M}_{m\times n}$, $\exists! \mathsf{T} \in \mathcal{L}(\mathsf{V}, \mathsf{W})$ such that $\Phi(\mathsf{T}) = A$.

Let $\beta = \{\beta_1, \beta_2, \ldots, \beta_n\}$, $\gamma = \{\gamma_1, \gamma_2, \ldots, \gamma_m\}$, and let $A = (a_{ij})$ be a given $m \times n$ matrix. Applying theorem 16 to the starting vector space V with the ordered basis β and a destination vector space W with the set of vectors:

$$\begin{aligned} w_1 &= a_{11}\gamma_1 + a_{21}\gamma_2 + \cdots + a_{m1}\gamma_m \\ w_2 &= a_{12}\gamma_1 + a_{22}\gamma_2 + \cdots + a_{m2}\gamma_m \\ &\vdots \\ w_n &= a_{1n}\gamma_1 + a_{2n}\gamma_2 + \cdots + a_{mn}\gamma_m \end{aligned}$$

where $\gamma = \{\gamma_1, \gamma_2, \ldots, \gamma_n\}$ form an ordered basis for W. Then there exists a unique linear transformation $\mathsf{T} : \mathsf{V} \mapsto \mathsf{W}$ such that:

$$\mathsf{T}(\beta_j) = w_j = a_{1j}\gamma_1 + a_{2j}\gamma_2 + \cdots + a_{mj}\gamma_m, \ j = 1, 2, \ldots, n.$$

Which implies that $[\mathsf{T}]_\beta^\gamma = A$, that is, $\Phi(\mathsf{T}) = A$. Therefore Φ is an isomorphism. □

Definition 18. Standard Representation. Let β be an ordered basis for an n-dimensional vector space V over the field F. The **standard representation** of V with respect to β is the function $\phi_\beta : \mathsf{V} \mapsto F^n$ defined by:

$$\phi_\beta(v) = [v]_\beta, \ \forall v \in \mathsf{V}.$$

Example 10.0.2. Let $\beta = \{(1,0), (0,1)\}$ and $\gamma = \{(1,2), (3,4)\}$ be ordered bases for $\mathbb{R}^2$. For $v = (1, -2)$, we have

$$\phi_\beta(v) = [v]_\beta = \begin{bmatrix} 1 \\ -2 \end{bmatrix} \quad \text{and} \quad \phi_\gamma(v) = [v]_\gamma = \begin{bmatrix} -5 \\ 2 \end{bmatrix}.$$

Theorem 28. For any finite-dimensional vector space V with ordered basis β, ϕ_β is an isomorphism.

Now let V and W be vector spaces of dimension n and m respectively, and let $\mathsf{T} \in \mathcal{L}(\mathsf{V}, \mathsf{W})$. Next, define $A = [\mathsf{T}]_\beta^\gamma$, where β and γ are arbitrary ordered bases of V and W respectively. Then the following **commutative diagram** maps the relationships between $\mathsf{V}, \mathsf{W}, F^n$, and F^m:

$$\begin{tikzcd} \mathsf{V} \arrow[r, "\mathsf{T}"] \arrow[d, "\phi_\beta"'] & \mathsf{W} \arrow[d, "\phi_\gamma"] \\ F^n \arrow[r, "\mathsf{L}_A"'] & F^m \end{tikzcd}$$

It also shows that there are two composites of linear transformations that map V into F^m, and thus we can conclude that

$$\mathsf{L}_A \phi_\beta = \phi_\gamma \mathsf{T},$$

that is, that the diagram "commutes". This diagram allows us to transfer operations on abstract vector spaces to ones on F^n and F^m.

Example 10.0.3. Let us define $\mathsf{T} \in \mathcal{L}(P_3(\mathbb{R}), P_2(\mathbb{R}))$ by:

$$\mathsf{T}(f(x)) = f'(x).$$

Let β and γ be the standard ordered bases for $P_3(\mathbb{R})$ and $P_2(\mathbb{R})$ respectively, and let $\phi_\beta : P_3(\mathbb{R}) \mapsto \mathbb{R}^4$ and $\phi_\gamma : P_2(\mathbb{R}) \mapsto \mathbb{R}^3$ be the corresponding standard representations of $P_3(\mathbb{R})$ and $P_2(\mathbb{R})$. If $A = [\mathsf{T}]_\beta^\gamma$ then

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

Now consider $p(x) = 2 + x - 3x^2 + 5x^3$, we can see that:

$$\mathsf{L}_A \phi_\beta(p(x)) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ -3 \\ 5 \end{bmatrix} = \begin{bmatrix} 1 \\ -6 \\ 15 \end{bmatrix}.$$

But since $\mathsf{T}(p(x)) = p'(x) = 1 - 6x + 15x^2$, we have:

$$\phi_\gamma \mathsf{T}(p(x)) = \begin{bmatrix} 1 \\ -6 \\ 15 \end{bmatrix}.$$

Lecture No. 11

The Change of Coordinate Matrix

In this lecture, we study how a coordinate vector relative to one basis can be changed into a coordinate vector relative to the other.
Example 11.0.1. Consider the equation

$$2x^2 - 4xy + 5y^2 = 1.$$

It is hard to see what curve this equation represents. If we make the following changes of variables:

$$\begin{aligned} x &= \frac{2}{\sqrt{5}}x' - \frac{1}{\sqrt{5}}y' \\ y &= \frac{1}{\sqrt{5}}x' + \frac{2}{\sqrt{5}}y' \end{aligned}$$

then we obtain the equation:

$$(x')^2 + 6(y')^2 = 1,$$

which is the equation of an ellipse. In this case, the coordinates of a point relative to the unit vectors of the standard ordered basis

$$\beta = \left\{ e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$$

is changed to new coordinates of the same point, but relative to the unit vectors of the new ordered basis:

$$\beta' = \left\{ \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{5}} \begin{bmatrix} -1 \\ 2 \end{bmatrix} \right\}.$$

Notice that the equations of the change of variables shown above can be represented in the form of a matrix equation:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix}.$$

Notice also that the matrix:

$$Q = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}$$

is equal to $[I]^{\beta}_{\beta'}$, where I is the identity transformation of $\mathbb{R}^2$. This phenomenon is true in general.

Theorem 29. Let V be a finite dimensional vector space with the ordered bases β and β'. Also let $Q = [I_{\mathsf{V}}]^{\beta}_{\beta'}$. Then Q is invertible, and for any $v \in \mathsf{V}$:

$$[v]_{\beta} = Q[v]_{\beta'} = [I_{\mathsf{V}}]^{\beta}_{\beta'}[v]_{\beta'}.$$

There is a mathematical name assigned to the matrix Q.

Definition 19. Change of Coordinate Matrix. For a given finite vector space V, that has ordered bases β and β', the matrix

$$Q = [I_{\mathsf{V}}]^{\beta}_{\beta'},$$

is called a **change of coordinate matrix**, which changes β'-coordinates into β-coordinates.

Example 11.0.2. Let the vector space in question be $\mathbb{R}^2$, with $\beta = \{(1,1),(1,-1)\}$ and $\beta' = \{(2,4),(3,1)\}$. Find the matrix that changes β'-coordinates into β-coordinates.

Solution. By definition, $Q = [I_{\mathsf{V}}]^{\beta}_{\beta'}$, so we have to represent I_{V} at β' in terms of β:

$$I_{\mathsf{V}} \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 1 \end{bmatrix} - 1 \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

and

$$I_{\mathsf{V}} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 1 \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Hence

$$Q = \begin{bmatrix} 3 & 2 \\ -1 & 1 \end{bmatrix}.$$

Theorem 30. Let T be a linear operator on a finite-dimensional vector space V. Also let β and β' be ordered bases for V, where Q is the change of coordinate matrix that changes β' into β. Then:

$$[\mathsf{T}]_{\beta'} = Q^{-1}[\mathsf{T}]_{\beta}Q.$$

Proof. Let $I := I_{\mathsf{V}}$ be the identity transformation on V. Then $\mathsf{T} = I \circ \mathsf{T} = \mathsf{T} \circ I$. Hence,

$$Q[\mathsf{T}]_{\beta'} = [I]^{\beta}_{\beta'}[\mathsf{T}]^{\beta'}_{\beta'} = [I \circ \mathsf{T}]^{\beta}_{\beta'} = [\mathsf{T} \circ I]^{\beta}_{\beta'} = [\mathsf{T}]^{\beta}_{\beta}[I]^{\beta}_{\beta'} = [\mathsf{T}]_{\beta}Q.$$

Therefore, $[\mathsf{T}]_{\beta'} = Q^{-1}[\mathsf{T}]_{\beta}Q$. $\square$

Example 11.0.3. Let T be the following linear operator on $\mathbb{R}^2$:

$$\mathsf{T}\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 3a - b \\ a - 3b \end{bmatrix}.$$

Also, let $\beta = \{(1,1),(1,-1)\}$ and $\beta' = \{(2,4),(3,1)\}$ as in example 11.0.2. Find $[\mathsf{T}]_\beta$, as well as the matrix Q that changes β'-coordinates into β-coordinates, and $[\mathsf{T}]_{\beta'}$, using the theorem above.
Solution. We have:

$$[\mathsf{T}]_\beta = \left[\left(\mathsf{T}\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)_\beta, \left(\mathsf{T}\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right)_\beta\right] = \begin{bmatrix} 3 & 1 \\ -1 & 3 \end{bmatrix}.$$

By example 11.0.2, we know what Q is, so we find its inverse:

$$Q = \begin{bmatrix} 3 & 2 \\ -1 & 1 \end{bmatrix} \implies Q^{-1} = \frac{1}{5}\begin{bmatrix} 1 & -2 \\ 1 & 3 \end{bmatrix}.$$

Hence,

$$\begin{aligned} [\mathsf{T}]_{\beta'} &= Q^{-1}[\mathsf{T}]_\beta Q \\ &= \frac{1}{5}\begin{bmatrix} 1 & -2 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 3 & 1 \\ -1 & 3 \end{bmatrix}\begin{bmatrix} 3 & 2 \\ -1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 4 & 1 \\ -2 & 2 \end{bmatrix} \end{aligned}$$

We have the following result that follows from the theorem.
Corollary 11.0.1. Let $A \in \mathsf{M}_{n\times n}(F)$ and let γ be an ordered basis for F^n. Then

$$[\mathsf{L}_A]_\gamma = Q^{-1}AQ,$$

where Q is the $n \times n$ matrix whose j-th column is the j-th vector of γ.

Proof. This is a special case of the theorem, where $\mathsf{T} = \mathsf{L}_A$, the left-multiplication operator on F^n. Indeed, in the conclusion of the theorem

$$[\mathsf{T}]_{\beta'} = Q^{-1}[\mathsf{T}]_\beta Q,$$

we take, for our case, γ instead of β', where β is the **standard** ordered basis for F^n, and we use L_A instead of T. Then we get

$$[\mathsf{L}_A]_\gamma = Q^{-1}[\mathsf{L}_A]_\beta Q.$$

Since β is the **standard** ordered basis for F^n, we have $[\mathsf{L}_A]_\beta = A$ and $Q = [I]_\gamma^\beta$, which is precisely the $n \times n$ matrix whose j-th column is the j-th vector of γ. □

Example 11.0.4. Let

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 3 \\ 0 & -1 & 0 \end{bmatrix},$$

and

$$\gamma = \left\{\begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right\}, \text{ an ordered basis for } \mathbb{R}^3.$$

Using the corollary, find $[\mathsf{L}_A]_\gamma$.

Solution. Let Q be then 3×3 matrix whose j-th column is the j-th vector of γ. Then

$$Q = \begin{bmatrix} -1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \qquad Q^{-1} = \begin{bmatrix} -1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Therefore,

$$\begin{aligned} [\mathsf{L}_A]_\gamma &= Q^{-1}AQ \\ &= \begin{bmatrix} -1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 3 \\ 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} -1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 2 & 8 \\ -1 & 4 & 6 \\ 0 & -1 & -1 \end{bmatrix}. \end{aligned}$$

We define the relationship between the matrices $[\mathsf{T}]_\beta$ and $[\mathsf{T}]_{\beta'}$.

Definition 20. Similarity. Let $A, B \in \mathsf{M}_{n\times n}(F)$. We say that B is **similar** to A if there exists an *invertible* matrix Q such that

$$B = Q^{-1}AQ,$$

denoted as $B \sim A$.

Lecture No. 12

Eigenvectors & Eigenvalues

The study of eigenvectors and eigenvalues is used to help solve the diagonalization problem.

Example 12.0.1. Find an expression of the reflection T about the line $y = 2x$.
Solution. We can use theorem 16. Note that the point $(1, 2)$ lies ont he straight line $y = 2x$ and the point $(-2, 1)$ lies on the line perpendicular to the first one. Then it is clear that

$$\mathsf{T}\begin{bmatrix}1\\2\end{bmatrix} = \begin{bmatrix}1\\2\end{bmatrix}, \qquad \mathsf{T}\begin{bmatrix}-2\\1\end{bmatrix} = -\begin{bmatrix}-2\\1\end{bmatrix} = \begin{bmatrix}2\\-1\end{bmatrix}.$$

Therefore, if we take

$$\beta' = \left\{\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}-2\\1\end{bmatrix}\right\}$$

as an ordered basis for $\mathbb{R}^2$, then we have

$$\mathsf{T}\begin{bmatrix}1\\2\end{bmatrix} = \begin{bmatrix}1\\2\end{bmatrix} = 1\begin{bmatrix}1\\2\end{bmatrix} + 0\begin{bmatrix}-2\\1\end{bmatrix}$$

and

$$\mathsf{T}\begin{bmatrix}-2\\1\end{bmatrix} = \begin{bmatrix}2\\-1\end{bmatrix} = 0\begin{bmatrix}1\\2\end{bmatrix} - 1\begin{bmatrix}-2\\1\end{bmatrix}.$$

That is,

$$[\mathsf{T}]_{\beta'} = \begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}.$$

Note. This matrix is **diagonal**.

Furthermore, if β is the standard basis for $\mathbb{R}^2$, then we can determine the matrix that changes β'-coordinates into β-coordinates, Q, and its inverse Q^{-1}:

$$Q = [I]_{\beta'}^{\beta} = \begin{bmatrix}1 & -2\\2 & 1\end{bmatrix}, \qquad Q^{-1} = \frac{1}{5}\begin{bmatrix}1 & 2\\-2 & 1\end{bmatrix}.$$

Then since $Q^{-1}[\mathsf{T}]_\beta Q = [\mathsf{T}]_{\beta'}$, we can solve the equation to get

$$[\mathsf{T}]_\beta = Q[\mathsf{T}]_{\beta'}Q^{-1} = \frac{1}{5}\begin{bmatrix}-3 & 4\\4 & 3\end{bmatrix}.$$