2: LINEAR TRANSFORMATIONS AND MATRICES

STEVEN HEILMAN

Contents

1. Review

Here are some theorems from last time that we will need below.

Corollary 1.1. *Let V be a vector space over a field $\mathbf{F}$. Assume that B is a finite basis of V, and B has exactly d elements. Then*

(f) *Any set of linearly independent elements of V is contained in a basis of V.*

Theorem 1.2. *Let V be a finite-dimensional vector space over a field $\mathbf{F}$. Let W be a subspace of V. Then W is also finite-dimensional, and $\dim(W) \leq \dim(V)$. Moreover, if $\dim(W) = \dim(V)$, then $W = V$.*

Theorem 1.3 (**Existence and Uniqueness of Basis Coefficients**). *Let $\{u_1, \ldots, u_n\}$ be a basis for a vector space V over a field $\mathbf{F}$. Then for any vector $u \in V$, there exist unique scalars $\alpha_1, \ldots, \alpha_n \in \mathbf{F}$ such that*

$$u = \sum_{i=1}^{n} \alpha_i u_i.$$

2. Linear Transformations

The general approach to the foundations of mathematics is to study certain spaces, and then to study functions between these spaces. In this course we follow this paradigm. Up until now, we have been studying properties of vector spaces. Vector spaces have a linear structure, and so it is natural to deal with functions between vector spaces that preserve this linear structure. That is, we will concern ourselves with linear transformations between vector spaces. For finite-dimensional spaces, it will turn out that linear transformations can be represented by the action of a matrix on a vector. However, for infinite-dimensional

Date: February 17, 2017.

spaces, this representation doesn't quite hold anymore. (Though, thinking by the way of analogy allows many results for infinite-dimensional linear transformations to nearly follow from the finite-dimensional case.) In any case, we can get a good deal of mileage by simply talking about abstract linear transformations, without addressing matrices at all. We will begin this approach below.

Definition 2.1. Let V and W be vector spaces over a field $\mathbf{F}$. We call a function $T\colon V \to W$ a **linear transformation** from V to W if, for all $v, v' \in V$ and for all $\alpha \in \mathbf{F}$,

(a) $T(v + v') = T(v) + T(v')$. (T preserves vector addition.)
(b) $T(\alpha v) = \alpha T(v)$. ($T$ preserves scalar multiplication.)

Exercise 2.2. Let $T\colon V \to W$ be a linear transformation. Show that $T(0) = 0$.

Example 2.3. Define $T(v) := 0$. Then T is linear. This T is known as the zero transformation.

Example 2.4. Define $T\colon V \to V$ by $T(v) := v$. Then T is linear.

Example 2.5. Define $T\colon \mathbf{R} \to \mathbf{R}$ by $T(x) := x^2$. Then T is **not** linear.

Example 2.6. Let $a, b, c, d \in \mathbf{R}$. Define $T\colon \mathbf{R}^2 \to \mathbf{R}^2$ by

$$T\begin{pmatrix} x \\ y \end{pmatrix} := \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}.$$

Then T is linear.

Example 2.7. Define $T\colon C^\infty(\mathbf{R}) \to C^\infty(\mathbf{R})$ by $T(f) := df/dt$. Then T is linear.

Example 2.8. Define $T\colon C^\infty(\mathbf{R}) \to C^\infty(\mathbf{R})$ by $T(f) := \int_0^1 f(t)dt$. Then T is linear.

Remark 2.9. The set $\mathcal{L}(V, W)$ of all linear transformations from $V \to W$ is itself a vector space over $\mathbf{F}$. We write $\mathcal{L}(V) := \mathcal{L}(V, V)$. Given linear transformations $S, T\colon V \to W$, we define $S + T$ so that, for all $v \in V$, $(S + T)(v) := S(v) + T(v)$. Also, for any $\alpha \in \mathbf{F}$, we define αT so that, for all $v \in V$, $(\alpha T)(v) := \alpha(T(v))$.

3. Null spaces, range, coordinate bases

Definition 3.1 (Null Space). Let V, W be vector spaces over a field $\mathbf{F}$. Let $T\colon V \to W$ be a linear transformation. The **null space** of T, denoted $N(T)$, is defined as

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

Remark 3.2. $N(T)$ is also referred to as the **kernel** of T. Note that $N(T)$ is a subspace of V, so its dimension can be defined.

Definition 3.3 (Nullity). Let V, W be vector spaces over a field $\mathbf{F}$. Let $T\colon V \to W$ be a linear transformation. The **nullity** of T, denoted $\text{nullity}(T)$, is defined as

$$\dim(N(T)).$$

Theorem 3.4. *Let V, W be vector spaces over a field $\mathbf{F}$. Let $T\colon V \to W$ be a linear transformation. Then T is injective if and only if $N(T) = \{0\}$.*

Proof. Suppose T is injective. We will show that $N(T) = \{0\}$. Note that $T(0) = 0$ by Exercise 2.2, so $\{0\} \subseteq N(T)$. It now suffices to show that $N(T)$ has only one element, which we prove by contradiction. Suppose there exist $v, v' \in N(T)$ such that $v \neq v'$. Since T is injective, $T(v) \neq T(v')$. But $v, v' \in N(T)$ imply $0 = T(v) = T(v')$, a contradiction. We conclude that $N(T)$ has only one element, as desired.

Now, suppose $N(T) = \{0\}$. We will show that T is injective. Let $v, v' \in V$ such that $T(v) = T(v')$. By linearity of T, $T(v - v') = T(v) - T(v') = 0$, so $v - v' \in N(T)$. Since $N(T) = \{0\}$, $v - v' = 0$, so that $v = v'$, proving the injectivity of T. □

Definition 3.5 (**Range**)**.** Let $T : V \to W$ be a linear transformation. The **range** of T, denoted $R(T)$, is defined as

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

Remark 3.6. Note that $R(T)$ is a subspace of W, so its dimension can be defined.

Definition 3.7 (**Rank**)**.** Let V, W be vector spaces over a field $\mathbf{F}$. Let $T : V \to W$ be a linear transformation. The **rank** of T, denoted $\mathrm{rank}(T)$, is defined as

$$\dim(R(T)).$$

Exercise 3.8. Let $T : V \to W$ be a linear transformation. Prove that $N(T)$ is a subspace of V and that $R(T)$ is a subspace of W.

Theorem 3.9 (**Dimension Theorem/ Rank-Nullity Theorem**)**.** *Let V, W be vector spaces over a field $\mathbf{F}$. Let $T : V \to W$ be linear. If V is finite-dimensional, then*

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

Proof. Since V is finite dimensional, and $N(T) \subseteq V$ is a subspace, $N(T)$ is finite dimensional by Theorem 1.2. In particular, a basis $\{v_1, \ldots, v_k\}$ for $N(T)$ exists, by the definition of finite-dimensionality. So, the set $\{v_1, \ldots, v_k\} \subseteq V$ is linearly independent. By Corollary 1.1(f), the set $\{v_1, \ldots, v_k\}$ is therefore contained in a basis for V. (Since V is finite-dimensional, a basis for V exists, so we can apply Corollary 1.1.) So, we have a basis $\{v_1, \ldots, v_k, u_1, \ldots, u_m\}$ for V. That is, $\mathrm{nullity}(T) = k$ and $\dim(V) = k + m$. It remains to show that $\mathrm{rank}(T) = m$.

We now show that $\mathrm{rank}(T) = m$. To show this, it suffices to show that $\{Tu_1, \ldots, Tu_m\}$ is a basis for $R(T)$. Let us therefore show that $\{Tu_1, \ldots, Tu_m\}$ is a linearly independent set. We prove this by contradiction. Suppose $\{Tu_1, \ldots, Tu_m\}$ is not a linearly independent set. Then there exist $\alpha_1, \ldots, \alpha_m \in \mathbf{F}$ which are not all equal to zero, such that

$$\sum_{i=1}^{m} \alpha_i T u_i = 0.$$

Since T is linear, we can rewrite this as

$$T\left(\sum_{i=1}^{m} \alpha_i u_i\right) = 0.$$

That is, $\sum_{i=1}^{m} \alpha_i u_i \in N(T)$. Since $\{v_1, \ldots, v_k\}$ is a basis for $N(T)$, there exist scalars $\beta_1, \ldots, \beta_k \in \mathbf{F}$ such that

$$\sum_{i=1}^{m} \alpha_i u_i = \sum_{i=1}^{k} \beta_i v_i.$$

That is,

$$\sum_{i=1}^{m} \alpha_i u_i - \sum_{i=1}^{k} \beta_i v_i = 0. \qquad (*)$$

Since the set $\{v_1, \ldots, v_k, u_1, \ldots, u_m\}$ is a basis for V, this set is linearly independent. So all the coefficients in $(*)$ are zero. In particular, $\alpha_1 = \cdots = \alpha_m = 0$. But we assumed that some α_i was nonzero. Since we have achieved a contradiction, we conclude that $\{Tu_1, \ldots, Tu_m\}$ is a linearly independent set.

It now remains to show that $\{Tu_1, \ldots, Tu_m\}$ is a spanning set of $R(T)$. Let $w \in R(T)$. We need to show that w is a linear combination of $\{Tu_1, \ldots, Tu_m\}$. Since $w \in R(T)$, there exists $u \in V$ such that $T(u) = w$. Since $\{v_1, \ldots, v_k, u_1, \ldots, u_m\}$ is a basis for V, there exist scalars $\gamma_1, \ldots, \gamma_k, \delta_1, \ldots, \delta_m \in \mathbf{F}$ such that $u = \sum_{i=1}^{k} \gamma_i v_i + \sum_{i=1}^{m} \delta_i u_i$. Applying T to both sides of this equation, and recalling that $v_i \in N(T)$ for all $i \in \{1, \ldots, k\}$, we get

$$T(u) = T\left(\sum_{i=1}^{k} \gamma_i v_i + \sum_{i=1}^{m} \delta_i u_i\right) = T\left(\sum_{i=1}^{m} \delta_i u_i\right) = \sum_{i=1}^{m} \delta_i T(u_i). \qquad (**)$$

Since $w = T(u)$, we have just expressed w as a linear combination of $\{Tu_1, \ldots, Tu_m\}$, as desired. We conclude that $\{Tu_1, \ldots, Tu_m\}$ is a spanning set for $R(T)$, so that $\mathrm{rank}(T) = m$, as desired. □

Lemma 3.10. *Let V and W be finite-dimensional vector spaces over a field $\mathbf{F}$. Assume that $\dim(V) = \dim(W)$. Let $T\colon V \to W$ be linear. Then T is one-to-one if and only if T is onto.*

Proof. We only prove the forward implication. Suppose T is one-to-one. Then $N(T) = \{0\}$ by Theorem 3.4. By the Dimension Theorem (Theorem 3.9), $\mathrm{rank}(T) = \dim(V)$. Since $\dim(V) = \dim(W)$, $\mathrm{rank}(T) = \dim(W)$. Since $R(T)$ is a subspace of W, and $\dim(R(T)) = \dim(W)$, we conclude that $R(T) = W$ by Theorem 1.2. So, T is onto, as desired. □

Exercise 3.11. Prove the reverse implication of Lemma 3.10.

Exercise 3.12. Define $T\colon C(\mathbf{R}) \to C(\mathbf{R})$ by $Tf(x) := \int_0^x f(t)dt$. Note that T is linear and one-to-one, but not onto, since there does not exist $f \in C(\mathbf{R})$ such that $T(f)(x) = 1$ for all $x \in \mathbf{R}$. Define $S\colon P(\mathbf{R}) \to P(\mathbf{R})$ by $Sf := df/dt$. Note that S is linear and onto, but S is not one-to-one, since S maps the constant function 1 to the zero function. How can you reconcile these facts with Lemma 3.10?

4. Linear Transformations and Bases

We will now isolate a few facts related to the main steps of the proof of the Dimension Theorem. These facts will be useful for us in our later discussion of isomorphism.

Theorem 4.1. *Let V, W be vector spaces over a field $\mathbf{F}$. Let $T\colon V \to W$ be a linear transformation. Assume that $\{v_1, \ldots, v_n\}$ spans V. Then $\{Tv_1, \ldots, Tv_n\}$ spans $R(T)$.*

Proof. Let $w \in R(T)$. We need to express w as a linear combination of $\{Tv_1, \ldots, Tv_n\}$. Since $w \in R(T)$, there exists $v \in V$ such that $T(v) = w$. Since $\{v_1, \ldots, v_n\}$ spans V, there exist scalars $\alpha_1, \ldots, \alpha_n \in \mathbf{F}$ such that $v = \sum_{i=1}^{n} \alpha_i v_i$. Applying T to both sides of this equality, and then using linearity of T,

$$T(v) = T\left(\sum_{i=1}^{n} \alpha_i v_i\right) = \sum_{i=1}^{n} \alpha_i T(v_i).$$

Since $w = T(v)$, we have expressed w as a linear combination of $\{Tv_1, \ldots, Tv_n\}$, as desired. $\square$

Theorem 4.2. *Let V, W be vector spaces over a field $\mathbf{F}$. Let $T\colon V \to W$ be a linear transformation which is one-to-one. Assume that $\{v_1, \ldots, v_n\}$ is linearly independent. Then $\{T(v_1), \ldots, T(v_n)\}$ is also linearly independent.*

Proof. We argue by contradiction. Assume that $\{T(v_1), \ldots, T(v_n)\}$ is linearly dependent. Then there exist scalars $\alpha_1, \ldots, \alpha_n \in \mathbf{F}$ not all equal to zero such that $\sum_{i=1}^{n} \alpha_i T(v_i) = 0$. Applying linearity of T, this equation says $T(\sum_{i=1}^{n} \alpha_i v_i) = 0$. Since T is one-to-one, we must have

$$\sum_{i=1}^{n} \alpha_i v_i = 0.$$

However, the set $\{v_1, \ldots, v_n\}$ is linearly independent, so we must have $\alpha_1 = \cdots = \alpha_n = 0$. But at least one α_i must be nonzero, a contradiction. We conclude that $\{T(v_1), \ldots, T(v_n)\}$ is linearly independent, as desired. $\square$

Corollary 4.3 (Bijections Preserve Bases). *Let V, W be vector spaces over a field $\mathbf{F}$. Let $T\colon V \to W$ be a linear transformation which is one-to-one and onto. Assume that $\{v_1, \ldots, v_n\}$ is a basis for V. Then $\{T(v_1), \ldots, T(v_n)\}$ is a basis for W. And therefore, $\dim(V) = \dim(W) = n$.*

Proof. Since $\{v_1, \ldots, v_n\}$ is a basis for V, $\{v_1, \ldots, v_n\}$ spans V. So, from Theorem 4.1, $\{T(v_1), \ldots, T(v_n)\}$ spans $R(T)$. Since T is onto, $R(T) = W$, so $\{T(v_1), \ldots, T(v_n)\}$ spans W. It remains to show that $\{T(v_1), \ldots, T(v_n)\}$ is linearly independent. Since $\{v_1, \ldots, v_n\}$ is a basis for V, $\{v_1, \ldots, v_n\}$ is linearly independent. So, from Theorem 4.2, $\{T(v_1), \ldots, T(v_n)\}$ is linearly independent, as desired. $\square$

As we now show, if $T\colon V \to W$ is linear and T is defined only on a basis of V, then this is sufficient to define T over all vectors in V. We phrase this theorem as a combined existence and uniqueness statement.

Theorem 4.4 (Rigidity of Linear Transformations). *Let V, W be vector spaces over a field $\mathbf{F}$. Assume that $\{v_1, \ldots, v_n\}$ is a basis for V. Let $\{w_1, \ldots, w_n\}$ be any vectors in W. Then there exists a unique linear transformation $T\colon V \to W$ such that $T(v_i) = w_i$ for all $i \in \{1, \ldots, n\}$.*

Proof. We first prove that T exists. Let $v \in V$. From Theorem 1.3, there exist unique scalars $\alpha_1, \ldots, \alpha_n \in \mathbf{F}$ such that $v = \sum_{i=1}^{n} \alpha_i v_i$. Suppose we define a map

$$T(\sum_{i=1}^{n} \alpha_i v_i) := \sum_{i=1}^{n} \alpha_i w_i. \qquad (*)$$

Observe that $T\colon V \to W$ is a map. In particular, since the scalars $\alpha_1, \ldots, \alpha_n \in \mathbf{F}$ depend uniquely on v, T is well-defined. We now check that $T(v_i) = w_i$ for all $i \in \{1, \ldots, n\}$. Note that

$$v_i = 1 \cdot v_i + \sum_{1 \le j \le n \colon j \ne i} 0 \cdot v_j.$$

So, plugging this formula into $(*)$ shows that $T(v_i) = w_i$.

We now need to verify that T is linear. Let $\alpha \in \mathbf{F}$. We first verify that $T(\alpha v) = \alpha T(v)$.

$$\begin{aligned} T(\alpha v) &= T\left(\sum_{i=1}^{n} (\alpha\alpha_i) v_i\right) = \sum_{i=1}^{n} \alpha\alpha_i w_i \qquad \text{, by } (*) \\ &= \alpha(\sum_{i=1}^{n} \alpha_i w_i) = \alpha T(v) \qquad \text{, by } (*). \end{aligned}$$

So, $T(\alpha v) = \alpha T(v)$ for all $v \in V$ and for all $\alpha \in \mathbf{F}$. Let $v' \in V$. We now verify that $T(v+v') = T(v)+T(v')$. There exist unique scalars $\beta_1, \ldots, \beta_n \in \mathbf{F}$ such that $v' = \sum_{i=1}^{n} \beta_i v_i$. We now check

$$\begin{aligned} T(v+v') &= T\left(\sum_{i=1}^{n} (\alpha_i + \beta_i) v_i\right) = \sum_{i=1}^{n} (\alpha_i + \beta_i) w_i \qquad \text{, by } (*) \\ &= \sum_{i=1}^{n} \alpha_i w_i + \sum_{i=1}^{n} \beta_i w_i = T(v) + T(v') \qquad \text{, by } (*). \end{aligned}$$

In conclusion, the map T defined by $(*)$ is in fact a linear transformation.

We now finish the proof by showing that T is unique. Suppose some other linear transformation $T' \colon V \to W$ satisfies $T(v_i) = w_i$ for all $i \in \{1, \ldots, n\}$. Then $(T - T')(v_i) = 0$ for all $i \in \{1, \ldots, n\}$. So, for any $v \in V$, once we write $v = \sum_{i=1}^{n} \alpha_i v_i$, we have by linearity of $(T - T')$

$$(T - T')(v) = (T - T')(\sum_{i=1}^{n} \alpha_i v_i) = \sum_{i=1}^{n} \alpha_i (T - T')(v_i) = 0.$$

That is, $T - T' = 0$, so $T = T'$, as desired. □

5. Matrix Representation, Matrix Multiplication

Definition 5.1 (**Ordered Basis**). Let V be a finite-dimensional vector space over a field $\mathbf{F}$. An **ordered basis** for V is a an ordered set $(v_1, \ldots, v_n)$ of elements of V such that $\{v_1, \ldots, v_n\}$ is a basis of V.

Example 5.2. One ordered basis for $\mathbf{R}^2$ is $((1,0),(0,1))$.

Definition 5.3 (**Coordinate Vector**). Let $\beta = (v_1, \ldots, v_n)$ be an ordered basis for V, and let $v \in V$. From Theorem 1.3, there exist unique scalars such that $v = \sum_{i=1}^{n} \alpha_i v_i$. The scalars $\alpha_1, \ldots, \alpha_n$ are referred to as the **coordinates** of v with respect to β. We then define the **coordinate vector** of v relative to β by

$$[v]^{\beta} := \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$

Example 5.4. Let $v := (3,4)$. If $\beta = ((1,0),(0,1))$. Then $v = 3(1,0) + 4(0,1)$, so

$$[v]^{\beta} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}.$$

If $\beta' = ((1,-1),(1,1))$, then $v = (-1/2)(1,-1) + (7/2)(1,1)$, so

$$[v]^{\beta'} = \begin{pmatrix} -1/2 \\ 7/2 \end{pmatrix}.$$

If $\beta'' = ((3,4),(0,1))$, then $v = 1(3,4) + 0(0,1)$

$$[v]^{\beta''} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Definition 5.5 (Matrix Representation). Let V, W be finite-dimensional vector spaces. Let $\beta = (v_1, \ldots, v_n)$ be an ordered basis for V, and let $\gamma = (w_1, \ldots, w_m)$ be an ordered basis for W. Let $T \colon V \to W$ be linear. Then, for each $j \in \{1, \ldots, n\}$, there exist unique scalars $a_{1j}, \ldots, a_{mj} \in \mathbf{F}$ by Theorem 1.3 such that

$$T(v_j) = \sum_{i=1}^{m} a_{ij} w_i.$$

We therefore define the **matrix representation** of T with respect to the bases β and γ by

$$[T]_{\beta}^{\gamma} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

Remark 5.6. Note that the j^{th} column of $[T]_{\beta}^{\gamma}$ is exactly $[T(v_j)]^{\gamma}$, so that

$$[T]_{\beta}^{\gamma} = ([T(v_1)]^{\gamma}, [T(v_2)]^{\gamma}, \ldots, [T(v_n)]^{\gamma}).$$

So, if we have an arbitrary $v \in V$, and we write v uniquely as $v = \sum_{j=1}^{n} b_j v_j$ where $b_1, \ldots, b_n \in \mathbf{F}$, then by linearity, $Tv = \sum_{j=1}^{n} b_j T(v_j)$. That is,

$$Tv = \sum_{j=1}^{n} \sum_{i=1}^{m} b_j a_{ij} w_i = \sum_{i=1}^{m} (\sum_{j=1}^{n} b_j a_{ij}) w_i$$

If we also express Tv in the basis γ, so that $Tv = \sum_{i=1}^{m} c_i w_i$ where $c_1, \ldots, c_m \in \mathbf{F}$, then we equate like terms to get $c_i = \sum_{j=1}^{n} b_j a_{ij}$ for all $1 \leq i \leq m$. In matrix form, this becomes

$$\begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

Or, using our ordered basis notation,

$$[Tv]^{\gamma} = [T]_{\beta}^{\gamma} [v]^{\beta}.$$

Remark 5.7. The important point here is that a linear transformation $T: V \to W$ has a meaning that does not depend on any ordered basis. However, when we view T from different perspectives (i.e. we examine $[T]_\beta^\gamma$ for different ordered bases β, γ), then T may look very different. One of the major goals of linear algebra is to take a T and view it from the "correct" basis, so that $[T]_\beta^\gamma$ takes a rather simple form, and therefore T becomes easier to understand. For example, if we could find ordered bases β, γ such that $[T]_\beta^\gamma$ becomes a diagonal matrix, then this would be really nice, since diagonal matrices are fairly easy to understand, and therefore we would better understand T. Unfortunately, we cannot always find bases such that $[T]_\beta^\gamma$ becomes diagonal, but in certain cases this can be done.

Remark 5.8. If we have a linear transformation $T: V \to W$, then specifying ordered bases β, γ gives a matrix representation $[T]_\beta^\gamma$. Conversely, if we have a matrix representation $[T]_\beta^\gamma$, then we know how T acts on an ordered basis. So, by Theorem 4.4, we can recover $T: V \to W$ from the matrix representation $[T]_\beta^\gamma$.

Example 5.9. Let $T: \mathbf{R}^2 \to \mathbf{R}^2$ be the linear transformation that takes any vector $(x, y) \in \mathbf{R}^2$ and rotates this vector counterclockwise around the origin by an angle $\pi/2$. Note that this description of T does not make use of any ordered basis. Let us find two different matrix representations of T. We first use $\beta = \gamma = ((1,0),(0,1))$. In this case, note that $T(1,0) = (0,1)$ and $T(0,1) = (-1,0)$. So, $T(1,0) = 0(1,0) + 1(0,1)$ and $T(0,1) = -1(1,0) + 0(0,1)$, and

$$[T]_\beta^\gamma = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

We will now find a matrix representation of T that is the identity matrix. Let $\beta := ((1,0),(0,1))$ and let $\gamma := ((0,1),(-1,0))$. Then $T(1,0) = 1(0,1) + 0(-1,0)$ and $T(0,1) = 0(0,1) + 1(-1,0)$, so

$$[T]_\beta^\gamma = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Recall that, in Remark 2.9, we noted that the set $\mathcal{L}(V, W)$ of all linear transformations from $V \to W$ is itself a vector space over $\mathbf{F}$. Given linear transformations $S, T: V \to W$, we defined $S + T$ so that, for all $v \in V$, $(S + T)(v) := S(v) + T(v)$. Also, for any $\alpha \in \mathbf{F}$, we defined αT so that, for all $v \in V$, $(\alpha T)(v) := \alpha(T(v))$. We can also define the product, or composition, or linear transformations as follows.

Definition 5.10 (Product/Composition). Let U, V, W be vector spaces over a field $\mathbf{F}$. Let $S: V \to W$ and let $T: U \to V$ be linear transformations. We define the **product** or **composition** $ST: U \to W$ by the formula

$$ST(u) := S(T(u)) \qquad \forall\, u \in U.$$

Exercise 5.11. Using the linearity of S and T, show that $ST: U \to W$ is a linear transformation.

Definition 5.12 (Matrix Multiplication). Let A be an $m \times \ell$ matrix, and let B be an $n \times m$ matrix. That is, A is a collection of scalars arranged into m rows and ℓ columns as follows

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1\ell} \\ A_{21} & A_{22} & \cdots & A_{2\ell} \\ \vdots & \vdots & \cdots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{m\ell} \end{pmatrix}.$$

Then the $n \times \ell$ matrix BA is defined, so that the (k, i) entry of BA is given by

$$(BA)_{ki} := \sum_{j=1}^{m} B_{kj} A_{ji}. \qquad 1 \leq k \leq n,\, 1 \leq i \leq \ell$$

Definition 5.13. The $n \times n$ identity matrix I_n is defined by

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}.$$

Note that the composition of two linear transformations evidently has a natural definition in Definition 5.10. On the other hand, matrix multiplication in Definition 5.12 may have appeared somewhat unnatural at first sight. So, perhaps surprisingly, we now show that the composition of linear transformations exactly defines the matrix multiplication to which we are accustomed. Put another way, the matrix multiplication in Definition 5.12 is a realization, in coordinates, of the composition of two linear transformations.

Theorem 5.14 (**Equivalence of Composition and Matrix Multiplication**). *Suppose U, V, W are vector spaces over a field $\mathbf{F}$. Let $S: V \to W$ and let $T: U \to V$ be linear transformations. Assume that U is ℓ-dimensional and it has an ordered basis $\alpha = (u_1, \ldots, u_\ell)$. Assume that V is m-dimensional and it has an ordered basis $\beta = (v_1, \ldots, v_m)$. Assume that W is n-dimensional and it has an ordered basis $\gamma = (w_1, \ldots, w_n)$. Then*

$$[ST]_\alpha^\gamma = [S]_\beta^\gamma [T]_\alpha^\beta.$$

Proof. We first apply Definition 5.5 to T. Then there exist scalars $\{a_{ji}\}_{1 \leq j \leq m, 1 \leq i \leq \ell}$ such that, for each $1 \leq i \leq \ell$,

$$T(u_i) = \sum_{j=1}^{m} a_{ji} v_j. \tag{1}$$

That is,

$$[T]_\alpha^\beta = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1\ell} \\ a_{21} & a_{22} & \cdots & a_{2\ell} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{m\ell} \end{pmatrix}. \tag{2}$$

We now apply Definition 5.5 to S. Then there exist scalars $\{b_{kj}\}_{1 \leq k \leq n, 1 \leq j \leq m}$ such that, for each $1 \leq j \leq m$,

$$S(v_j) = \sum_{k=1}^{n} b_{kj} w_k. \tag{3}$$

That is,

$$[S]_\beta^\gamma = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \vdots & \vdots & \cdots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nm} \end{pmatrix}. \tag{4}$$

Applying S to both sides of (1) and using linearity of S,

$$\begin{aligned} S(T(u_i)) = S\left(\sum_{j=1}^{m} a_{ji} v_j\right) &= \sum_{j=1}^{m} a_{ji} S(v_j) \\ &= \sum_{j=1}^{m} a_{ji} \sum_{k=1}^{n} b_{kj} w_k \qquad \text{, by (3).} \end{aligned}$$

Changing the order of summation, we get

$$ST(u_i) = \sum_{k=1}^{n} \left(\sum_{j=1}^{m} b_{kj} a_{ji}\right) w_k. \tag{5}$$

So, for each $1 \leq k \leq n$ and $1 \leq i \leq \ell$, define

$$c_{ki} := \sum_{j=1}^{m} b_{kj} a_{ji}. \tag{6}$$

Then (5) becomes

$$ST(u_i) = \sum_{k=1}^{n} c_{ki} w_k. \tag{7}$$

That is, using the definitions of α and γ,

$$[ST]_\alpha^\gamma = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1\ell} \\ c_{21} & c_{22} & \cdots & c_{2\ell} \\ \vdots & \vdots & \cdots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{n\ell} \end{pmatrix}. \tag{8}$$

Finally, we use (2) and (4), and then perform the matrix multiplication

$$[S]_\beta^\gamma [T]_\alpha^\beta = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \vdots & \vdots & \cdots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nm} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1\ell} \\ a_{21} & a_{22} & \cdots & a_{2\ell} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{m\ell} \end{pmatrix}. \tag{9}$$

Then the matrix multiplication in (9), defined in Definition 5.12, agrees with the matrix in (8), because of (6). That is, $[ST]_\alpha^\gamma = [S]_\beta^\gamma [T]_\alpha^\beta$, as desired. □

5.1. **Matrices as Linear Transformations.** We showed in Theorem 5.14 that composing two linear transformations is equivalent to using matrix multiplication. We now belabor this point by beginning with a matrix, and then using the theory of linear transformations to prove associativity of matrix multiplication. We could prove that matrix multiplication is associative by taking three matrices and then writing out all the relevant terms. However, the "coordinate-free" approach below ends up being a bit more elegant. This proof strategy is part of a larger paradigm, in which "coordinate-free" proofs end up being more enlightening that coordinate-reliant proofs.

Definition 5.15. Consider the vector space $\mathbf{F}^n$ over the field $\mathbf{F}$. The **standard basis** for $\mathbf{F}^n$ is defined as

$$(e_1, \ldots, e_n) = ((1, 0, 0, \ldots, 0), (0, 1, 0, \ldots, 0), \ldots, (0, \ldots, 0, 1)).$$

Definition 5.16. Let A be an $m \times n$ matrix of scalars in a field $\mathbf{F}$. Define $L_A \colon \mathbf{F}^n \to \mathbf{F}^m$ by the formula

$$L_A(u) := Au, \quad \forall\, u \in \mathbf{F}^n.$$

Here we think of vectors in $\mathbf{F}^n$ and $\mathbf{F}^m$ as column vectors. Note that L_A is linear.

Lemma 5.17. *Let α be the standard basis of $\mathbf{F}^n$ and let β be the standard basis of $\mathbf{F}^m$. Let $A \in M_{m\times n}(\mathbf{F})$. Then $[L_A]_\alpha^\beta = A$. Let $T \colon \mathbf{F}^n \to \mathbf{F}^m$ be a linear transformation. Then $L_{[T]_\alpha^\beta} = T$.*

Proof. Let $u \in \mathbf{F}^n$ be a column vector. That is, there exist $\alpha_1, \ldots, \alpha_n \in \mathbf{F}^n$ such that

$$u = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}.$$

That is, $u = \sum_{i=1}^{n} \alpha_i u_i$. That is,

$$u = [u]^\alpha, \quad \forall\, u \in \mathbf{F}^n. \qquad (*)$$

Similarly,

$$v = [v]^\beta, \quad \forall\, v \in \mathbf{F}^m. \qquad (**)$$

From Remark 5.6,

$$[L_A(u)]^\beta = [L_A]_\alpha^\beta [u]^\alpha.$$

Applying $(*)$ and $(**)$, we get

$$L_A(u) = [L_A]_\alpha^\beta u.$$

Since $L_A(u) = Au$, we get

$$Au = [L_A]_\alpha^\beta u. \qquad \forall\, u \in \mathbf{F}^n$$

Using $u = e_i$ for any $i \in \{1, \ldots, n\}$ shows that the i^{th} column of A is equal to the i^{th} column of $[L_A]_\alpha^\beta$. So, $[L_A]_\alpha^\beta = A$, as desired.

Now, let $T \colon \mathbf{F}^n \to \mathbf{F}^m$ be a linear transformation. From Remark 5.6, for any $u \in \mathbf{F}^n$,

$$[T(u)]^\beta = [T]_\alpha^\beta [u]^\alpha.$$

Applying $(*)$ and $(**)$,

$$T(u) = [T]_\alpha^\beta u = L_{[T]_\alpha^\beta}(u). \qquad \forall\, u \in \mathbf{F}^n.$$

Therefore, $T = L_{[T]_\alpha^\beta}$, as desired. □

Lemma 5.18. *Let U, V, W, X be vector spaces over a field $\mathbf{F}$. Let $T: U \to V$, $S: V \to W$, $R: W \to X$ be three linear transformations. Then $R(ST) = (RS)T$.*

Proof. We are required to show that, for all $u \in U$, $R(ST)(u) = (RS)T(u)$. We repeatedly apply Definition 5.10 as follows.

$$R(ST)(u) = R(ST(u)) = R(S(T(u))) = RS(T(u)) = (RS)T(u).$$

□

Note that Lemma 5.18 was proven in a coordinate-free manner. We now combine Lemmas 5.17 and 5.18 to prove associativity of matrix multiplication, a statement that uses coordinates.

Corollary 5.19. *Let A be an $m \times \ell$ matrix, let B be an $n \times m$ matrix, and let C be a $k \times n$ matrix. Then $C(BA) = (CB)A$.*

Proof. From Lemma 5.18,

$$L_C(L_B L_A) = (L_C L_B) L_A. \tag{10}$$

Let $\alpha, \beta, \gamma, \delta$ be the standard bases for $\mathbf{F}^\ell, \mathbf{F}^m, \mathbf{F}^n$ and $\mathbf{F}^k$, respectively. Applying Theorem 5.14 twice to the left side of (10),

$$\begin{aligned} [L_C(L_B L_A)]_\alpha^\delta &= [L_C]_\gamma^\delta [L_B L_A]_\alpha^\gamma = [L_C]_\gamma^\delta ([L_B]_\beta^\gamma [L_A]_\alpha^\beta) \\ &= C(BA) \qquad \text{by Lemma 5.17.} \end{aligned} \tag{11}$$

Applying Theorem 5.14 twice to the right side of (10),

$$\begin{aligned} [(L_C L_B) L_A]_\alpha^\delta &= [L_C L_B]_\beta^\delta [L_A]_\alpha^\beta = ([L_C]_\gamma^\delta [L_B]_\beta^\gamma) [L_A]_\alpha^\beta \\ &= (CB)A \qquad \text{by Lemma 5.17.} \end{aligned} \tag{12}$$

Combining (10), (11) and (12) completes the proof. □

The following facts are proven in a similar manner.

Remark 5.20. Let A be an $m \times \ell$ matrix, let B be an $n \times m$ matrix. Then $L_B L_A = L_{BA}$.

Proof. Let α, β, γ be the standard bases for $\mathbf{F}^\ell, \mathbf{F}^m$ and $\mathbf{F}^n$ respectively. Applying Theorem 5.14 then Lemma 5.17,

$$[L_B L_A]_\alpha^\gamma = [L_B]_\beta^\gamma [L_A]_\alpha^\beta = BA.$$

Taking L of both sides and applying Lemma 5.17 to the left side shows that $L_B L_A = L_{BA}$. □

Remark 5.21. Let A be an $n \times m$ matrix, let B be an $n \times m$ matrix. Then $L_{A+B} = L_A + L_B$.

Proof. Let α, β be the standard bases for $\mathbf{F}^m$ and $\mathbf{F}^n$, respectively. Applying Lemma 5.17,

$$[L_A + L_B]_\alpha^\beta = [L_A]_\alpha^\beta + [L_B]_\alpha^\beta = A + B.$$

Taking L of both sides and applying Lemma 5.17 to the left side shows that $L_A + L_B = L_{A+B}$. □

6. Invertibility, Isomorphism

We now introduce the concept of invertibility. As will become clear, the invertibility of a linear transformation is closely related to our ability to find a "nice" matrix representation of the linear transformation.

Definition 6.1 (**Inverse**)**.** Let V, W be vector spaces over a field $\mathbf{F}$. Let $T\colon V \to W$ be a linear transformation. We say that a linear transformation $S\colon W \to V$ is the **inverse** of T if $TS = I_W$ and $ST = I_V$. We say that T is **invertible** if T has an inverse, and we denote the inverse by T^{-1}, so that $TT^{-1} = I_W$ and $T^{-1}T = I_V$.

Remark 6.2. If T is the inverse of S, then S is the inverse of T.

If an inverse of T exists, then it is unique, as we now show.

Lemma 6.3. *Let V, W be vector spaces over a field $\mathbf{F}$. Let $T\colon V \to W$ be a linear transformation. Let $S\colon W \to V$ be an inverse of T, and let $S'\colon W \to V$ be an inverse of T. Then $S = S'$.*

Proof. Using the definition of inverse,

$$S = SI_W = S(TS') = (ST)S' = I_V S' = S'.$$

□

Lemma 6.4. *Let V, W be vector spaces over a field $\mathbf{F}$. Let $T\colon V \to W$ be a linear transformation. If T has an inverse $S\colon W \to V$, then T must be one-to-one and onto.*

Proof. We first show that T is one-to-one. Suppose $v, v' \in V$ satisfy $T(v) = T(v')$. Applying S to both sides, $ST(v) = ST(v')$. That is, $v = v'$, so T is one-to-one, as desired.

We now show that T is onto. Let $w \in W$. We need to find $v \in V$ such that $T(v) = w$. Define $v := Sw$. Then $T(v) = TS(w) = w$, as desired. □

Example 6.5. The zero transformation $T\colon \mathbf{R}^2 \to \mathbf{R}^2$ defined by $T = 0$ is not onto, so T is not invertible.

We now prove the converse of Lemma 6.4

Lemma 6.6. *Let V, W be vector spaces over a field $\mathbf{F}$. Let $T\colon V \to W$ be a linear transformation. Suppose T is one-to-one and onto. Then there exists a linear transformation $S\colon W \to V$ that is the inverse of T.*

Proof. We first have to somehow define a linear transformation $S\colon W \to V$ that inverts T. Given any $w \in W$, since T is bijective, there exists a unique $v \in V$ such that $w = T(v)$. So, define

$$S(w) := v. \qquad (*)$$

Since v uniquely depends on w, the map $S\colon W \to V$ defined in this way is well-defined. We now show that S is linear. Let $w, w' \in W$. Since T is bijective, there exist unique $v, v' \in V$ such that $T(v) = w$ and $T(v') = w'$. In particular, by the definition $(*)$, $S(w) = v$ and $S(w') = v'$. Since T is linear, $T(v + v') = w + w'$. So, by the definition $(*)$, we have $S(w + w') = v + v' = S(w) + S(w')$. Now, let $\alpha \in \mathbf{F}$. Since $T(v) = w$ and T is linear, $T(\alpha v) = \alpha T(v) = \alpha w$. By the definition $(*)$, $S(\alpha w) = \alpha v$. Since $v = S(w)$, we therefore have $S(\alpha w) = \alpha S(w)$, as desired. So, S is linear.

It remains to show that S inverts T. Applying T to both sides of $(*)$, note that $TS(w) = T(v) = w$, so $TS = I_W$. Also, substituting $w = T(v)$ into $(*)$, we get $S(T(v)) = v$, so that $ST = I_V$, as desired. □

Combining Lemmas 6.4 and 6.6, we see that a linear transformation $T\colon V \to W$ is invertible if and only if T is one-to-one and onto. Invertible linear transformations are also known as **isomorphisms**.

Definition 6.7 (**Isomorphism**). Two vector spaces V, W over a field $\mathbf{F}$ are said to be **isomorphic** if there exists an invertible linear transformation $T\colon V \to W$ from one space to the other.

The notion of isomorphism allows us to reason about two vector spaces being the same (if they are isomorphic) or not the same (if they are not isomorphic). Many parts of mathematics, or science more generally, are concerned with classifying things according to whether they are they same or not the same. Within the context of vector spaces, this notion of isomorphism is most appropriate, since it asks for the linear structure of the vector space to be preserved. Within other mathematical contexts, different notions of isomorphism appear, though they all generally ask for the structures at hand to be preserved by a certain map.

Lemma 6.8. *Two finite-dimensional vectors spaces V, W over a field $\mathbf{F}$ are isomorphic if and only if* $\dim(V) = \dim(W)$.

Proof. Suppose V, W are isomorphic. Then there exists an invertible linear transformation $T\colon V \to W$. By Lemma 6.4, T is one-to-one and onto. In particular, $\text{nullity}(T) = 0$. By the Dimension Theorem (Theorem 3.9), $\text{rank}(T) = \dim(V)$. Since T is onto, $\text{rank}(T) = \dim(W)$. Therefore, $\dim(V) = \dim(W)$, as desired.

We now prove the reverse implication. Assume that $\dim(V) = \dim(W) = n$ for some $n \in \mathbf{N}$. Let $\{v_1, \ldots, v_n\}$ be a basis for V, and let $\{w_1, \ldots, w_n\}$ be a basis for W. By Theorem 4.4, there exists a linear transformation $T\colon V \to W$ such that $T(v_i) = w_i$ for all $i \in \{1, \ldots, n\}$. By Theorem 4.1, $\{w_1, \ldots, w_n\}$ spans $R(T)$. Since $\{w_1, \ldots, w_n\}$ also spans W, we have $R(T) = W$, so that T is onto. By Lemma 3.10 (using $\dim(V) = \dim(W)$), T is also one-to-one. So, T is an isomorphism, and V, W are isomorphic, as desired. □

Remark 6.9. If V has an ordered basis $\beta = (v_1, \ldots, v_n)$, then the coordinate map $\phi_\beta\colon V \to \mathbf{F}^n$ defined by

$$\phi_\beta(v) := [v]^\beta$$

is a linear transformation. It is also an isomorphism. Note that ϕ_β is one-to-one by Theorem 1.3, and ϕ_β is onto since, if we are given the coordinate vector $[v]^\beta = (\alpha_1, \ldots, \alpha_n)$, then $\phi_\beta(\sum_{i=1}^n \alpha_i v_i) = [v]^\beta$. So, ϕ_β is an isomorphism by Lemma 6.6. The book calls ϕ_β the **standard representation** of V with respect to β.

If we only care about linear transformations for finite-dimensional vector spaces over $\mathbf{R}$, then Lemma 6.8 and Theorem 5.14 show that it suffices to discuss real matrices and the vector spaces $\mathbf{R}^n$, $n \in \mathbf{N}$. However, our effort in developing the theory of linear transformations was not a waste of time. For example, the notion of isomorphism from Definition 6.7 is not very meaningful for infinite-dimensional vector spaces. For another example, when we introduce norms and inner products, the notion of isomorphism from Definition 6.7 becomes less meaningful, and finer properties of linear transformations become more relevant. Nevertheless, we will mostly discuss real matrices and $\mathbf{R}^n$ for the rest of the course.

6.1. **Invertibility and Matrices.**

Definition 6.10 (Inverse Matrix). Let A be an $m \times n$ matrix. We say that A has an **inverse** B if B is an $n \times m$ matrix such that $AB = I_m$ and such that $BA = I_n$. If A has an inverse, we say that A is an **invertible matrix**, and we write $B = A^{-1}$.

We now continue to emphasize the relation between linear transformations and matrices, as in Theorem 5.14 and Remark 5.6.

Theorem 6.11. *Let V, W be vector spaces over a field $\mathbf{F}$. Assume that α is an ordered basis for V with n elements, and assume that β is an ordered basis for W with m elements. Then a linear transformation $T\colon V \to W$ is invertible if and only if $[T]_\alpha^\beta$ is invertible. Also, $[T^{-1}]_\beta^\alpha = ([T]_\alpha^\beta)^{-1}$*

Proof. Suppose $T\colon V \to W$ has an inverse $T^{-1}\colon W \to V$. Then $TT^{-1} = I_W$ and $T^{-1}T = I_V$. So, applying Theorem 5.14,

$$[T]_\alpha^\beta [T^{-1}]_\beta^\alpha = [TT^{-1}]_\beta^\beta = [I_W]_\beta^\beta = I_m.$$

$$[T^{-1}]_\beta^\alpha [T]_\alpha^\beta = [T^{-1}T]_\alpha^\alpha = [I_V]_\alpha^\alpha = I_n.$$

So, $[T^{-1}]_\beta^\alpha$ is the inverse of $[T]_\alpha^\beta$, so that $[T]_\alpha^\beta$ is an invertible matrix.

We now prove the reverse implication. Suppose $[T]_\alpha^\beta$ is invertible. Then there exists an $n \times m$ matrix B such that $B[T]_\alpha^\beta = I_n$ and $[T]_\alpha^\beta B = I_m$. Write $\alpha = (v_1, \ldots, v_n)$, $\beta = (w_1, \ldots, w_m)$. We would like to have a linear transformation $S\colon W \to V$ such that $S(w_i) = \sum_{k=1}^n B_{ki} v_k$ for all $i \in \{1, \ldots, m\}$. If such an S exists, then $[S]_\beta^\alpha = B$. Such a linear transformation exists by Theorem 4.4. Therefore,

$$[I_V]_\alpha^\alpha = I_n = B[T]_\alpha^\beta = [S]_\beta^\alpha [T]_\alpha^\beta = [ST]_\alpha^\alpha.$$

$$[I_W]_\beta^\beta = I_m = [T]_\alpha^\beta B = [T]_\alpha^\beta [S]_\beta^\alpha = [TS]_\beta^\beta.$$

So, T is invertible, as desired. □

Corollary 6.12. *An $m \times n$ matrix A is invertible if and only if the linear transformation $L_A\colon \mathbf{F}^n \to \mathbf{F}^m$ is invertible. Also, $(L_A)^{-1} = L_{A^{-1}}$.*

Proof. Let α be the standard basis for $\mathbf{F}^n$ and let β be the standard basis for $\mathbf{F}^m$. Then

$$[L_A]_\alpha^\beta = A. \qquad (*)$$

So, by Theorem 6.11, L_A is invertible if and only if A is invertible. Also, from Theorem 6.11,

$$[L_A^{-1}]_\beta^\alpha = ([L_A]_\alpha^\beta)^{-1} = A^{-1} = [L_{A^{-1}}]_\beta^\alpha \qquad \text{, by } (*).$$

Therefore, $L_A^{-1} = L_{A^{-1}}$. □

Corollary 6.13. *Let A be an $m \times n$ matrix. If A is invertible, then $m = n$.*

Proof. Apply Corollary 6.12 and Lemma 6.8. □

Unfortunately, not all matrices are invertible. For example, the zero matrix is not invertible.

7. Change of Coordinates

Suppose we have two finite ordered bases β, β' for the same vector space V. Let $v \in V$. We would like a way to relate $[v]^{\beta}$ to $[v]^{\beta'}$. Using Remark 5.6 and that $I_V v = v$ for all $v \in V$, we have

$$[I_V]_{\beta}^{\beta'}[v]^{\beta} = [v]^{\beta'}.$$

That is, to relate $[v]^{\beta}$ to $[v]^{\beta'}$, it suffices to compute $[I_V]_{\beta}^{\beta'}$

Example 7.1. Let $\beta = ((2,0),(1,-1))$, and let $\beta' = ((0,1),(2,1))$ be two ordered bases of $\mathbf{R}^2$. Then

$$I_V(2,0) = (2,0) = -1(0,1) + 1(2,1).$$

$$I_V(1,-1) = (1,-1) = -(3/2)(0,1) + (1/2)(2,1).$$

So

$$[I_V]_{\beta}^{\beta'} = \begin{pmatrix} -1 & -3/2 \\ 1 & 1/2 \end{pmatrix}.$$

So, we can verify that $[I_V]_{\beta}^{\beta'}[v]^{\beta} = [v]^{\beta'}$. For example, choosing $v = (3,2)$, note that

$$[v]^{\beta} = \begin{pmatrix} 5/2 \\ -2 \end{pmatrix}, \qquad [v]^{\beta'} = \begin{pmatrix} 1/2 \\ 3/2 \end{pmatrix}, \qquad \begin{pmatrix} -1 & -3/2 \\ 1 & 1/2 \end{pmatrix}\begin{pmatrix} 5/2 \\ -2 \end{pmatrix} = \begin{pmatrix} 1/2 \\ 3/2 \end{pmatrix}.$$

Similarly, note that $[I_V]_{\beta'}^{\beta}$ is the inverse of $[I_V]_{\beta}^{\beta'}$, so

$$[I_V]_{\beta'}^{\beta} = \begin{pmatrix} 1/2 & 3/2 \\ -1 & -1 \end{pmatrix}.$$

Exercise 7.2. Show that $[I_V]_{\beta}^{\beta'}$ is invertible, with inverse $[I_V]_{\beta'}^{\beta}$.

Lemma 7.3. *Let V be a finite-dimensional vector space over a field $\mathbf{F}$. Let β, β' be two bases for V. Let $T\colon V \to V$ be a linear transformation. Define $Q := [I_V]_{\beta}^{\beta'}$. (From Theorem 6.11, Q is invertible.) Then $[T]_{\beta}^{\beta}$ and $[T]_{\beta'}^{\beta'}$ satisfy the following relation*

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

Proof. We first write $T = I_V T I_V$. Taking the matrix representation of both sides and then applying Theorem 5.14,

$$[T]_{\beta'}^{\beta'} = [I_V T I_V]_{\beta'}^{\beta'} = [I_V]_{\beta}^{\beta'}[T I_V]_{\beta'}^{\beta} = [I_V]_{\beta}^{\beta'}[T]_{\beta}^{\beta}[I_V]_{\beta'}^{\beta}.$$

From Theorem 6.11, $[I_V]_{\beta'}^{\beta} = ([I_V]_{\beta}^{\beta'})^{-1}$, completing the proof. □

Definition 7.4 (Similarity). Two $n \times n$ matrices A, B are said to be **similar** if there exists an invertible $n \times n$ matrix Q such that $A = QBQ^{-1}$.

Remark 7.5. In the context of Lemma 7.3, $[T]_{\beta'}^{\beta'}$ is similar to $[T]_{\beta}^{\beta}$.

8. Appendix: Notation

Let A, B be sets in a space X. Let m, n be a nonnegative integers. Let $\mathbf{F}$ be a field.

$\mathbf{Z} := \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$, the integers

$\mathbf{N} := \{0, 1, 2, 3, 4, 5, \ldots\}$, the natural numbers

$\mathbf{Q} := \{m/n \colon m, n \in \mathbf{Z}, n \neq 0\}$, the rationals

$\mathbf{R}$ denotes the set of real numbers

$\mathbf{C} := \{x + y\sqrt{-1} : x, y \in \mathbf{R}\}$, the complex numbers

$\emptyset$ denotes the empty set, the set consisting of zero elements

$\in$ means "is an element of." For example, $2 \in \mathbf{Z}$ is read as "2 is an element of $\mathbf{Z}$."

$\forall$ means "for all"

$\exists$ means "there exists"

$\mathbf{F}^n := \{(x_1, \ldots, x_n) \colon x_i \in \mathbf{F}, \forall\, i \in \{1, \ldots, n\}\}$

$A \subseteq B$ means $\forall\, a \in A$, we have $a \in B$, so A is contained in B

$A \smallsetminus B := \{x \in A \colon x \notin B\}$

$A^c := X \smallsetminus A$, the complement of A

$A \cap B$ denotes the intersection of A and B

$A \cup B$ denotes the union of A and B

$C(\mathbf{R})$ denotes the set of all continuous functions from $\mathbf{R}$ to $\mathbf{R}$

$P_n(\mathbf{R})$ denotes the set of all real polynomials in one real variable of degree at most n

$P(\mathbf{R})$ denotes the set of all real polynomials in one real variable

$M_{m\times n}(\mathbf{F})$ denotes the vector space of $m \times n$ matrices over the field $\mathbf{F}$

I_n denotes the $n \times n$ identity matrix

8.1. Set Theory.

Let V, W be sets, and let $f \colon V, W$ be a function. Let $X \subseteq V$, $Y \subseteq W$.

$$f(X) := \{f(v) \colon v \in V\}.$$

$$f^{-1}(Y) := \{v \in V \colon f(v) \in Y\}.$$

The function $f \colon V \to W$ is said to be **injective** (or **one-to-one**) if: for every $v, v' \in V$, if $f(v) = f(v')$, then $v = v'$.

The function $f \colon V \to W$ is said to be **surjective** (or **onto**) if: for every $w \in W$, there exists $v \in V$ such that $f(v) = w$.

The function $f \colon V \to W$ is said to be **bijective** (or a **one-to-one correspondence**) if: for every $w \in W$, there exists exactly one $v \in V$ such that $f(v) = w$. A function $f \colon V \to W$ is bijective if and only if it is both injective and surjective.

Two sets X, Y are said to have the same **cardinality** if there exists a bijection from V onto W.

The **identity map** $I \colon X \to X$ is defined by $I(x) = x$ for all $x \in X$. To emphasize that the domain and range are both X, we sometimes write I_X for the identity map on X.

Let V, W be vector spaces over a field $\mathbf{F}$. Then $\mathcal{L}(V, W)$ denotes the set of linear transformations from V to W, and $\mathcal{L}(V)$ denotes the set of linear transformations from V to V.

UCLA Department of Mathematics, Los Angeles, CA 90095-1555
E-mail address: heilman@math.ucla.edu