Matrices as Linear Transformations of Finite Dimensional Vector Spaces

In the previous section, we investigated an interesting example of a linear transformation: to matrix

$$M = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}$$

we associated the linear transformation T_M defined by

$$T_M\left(\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}\right) = M\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}.$$

This idea of a matrix as a linear transformation is one that we will discuss in great detail in this section. We begin with a lemma that you may have already guessed at:

Lemma. If A is an $m \times n$ matrix with entries in $\mathbb{F}$, then the function $T_A : \mathbb{F}^n \to \mathbb{F}^m$ defined by

$$T_A(u) = Au$$

is a linear transformation.

We will not prove the lemma, as it is clear from the properties of matrix operations. Instead, we will look at a generalization of this lemma to linear transformations between *any* vector spaces.

In fact, we wish to prove something more, a deeper result that will be integral to our understanding of vector spaces: *every linear transformation between finite dimensional vector spaces has a realization in terms of matrices and matrix multiplication.*

Coordinatization and Linear Transformations

First, we recall the definition for *coordinates of a vector*:

Definition. If $S = (v_1, v_2, \ldots, v_n)$ is a basis for the vector space V and the vector $v \in V$ is the linear combination

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n,$$

then the scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$ are called the *coordinates* of v relative to the basis S, and the vector

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$$

in $\mathbb{F}^n$ is called the *coordinate vector* of v relative to the basis S, denoted by

$$(v)_S = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}.$$

Again, coordinatization allows us to think of vectors in abstract spaces as *vectors in* $\mathbb{F}^n$. With this idea in mind, you may be able to see how we can think of matrices as linear transformations among abstract spaces. We illustrate this idea in the following example.

Example. The vectors $f_1(x) = 2x - 3$, $f_2(x) = x^2 + 1$, and $f_3(x) = 2x^2 - x$ are linearly independent and span the vector space $\mathcal{P}_2(\mathbb{R})$ of all real-valued polynomials of degree no more than 2, so that the list

$$B = (f_1,\ f_2,\ f_3)$$

is a basis for $\mathcal{P}_2(\mathbb{R})$. Let $p \in \mathcal{P}_2(\mathbb{R})$ with coordinates

$$(p)_B = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix},$$

and let I denote the 3×3 identity matrix. Define the function $T_I : \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}^3$ by

$$T_I(p) := I(p)_B.$$

1. Find the image of vector $p(x) = -x^2 - 5x + 10$ in $\mathbb{R}^3$ under the transformation.

2. Show that T_I is a linear transformation from $\mathcal{P}_2(\mathbb{R})$ to $\mathbb{R}^3$.

1. In order to find the image of the vector, we must first locate its coordinates under basis B. We look for the scalars α, β, γ so that

$$\alpha f_1 + \beta f_2 + \gamma f_3 = p.$$

Per usual, we think of the associated linear system for the equation above, whose augmented matrix is given by

$$\left(\begin{array}{ccc|c} 0 & 1 & 2 & -1 \\ 2 & 0 & -1 & -5 \\ -3 & 1 & 0 & 10 \end{array}\right);$$

row reducing, we have

$$\begin{aligned}
\left(\begin{array}{ccc|c} 0 & 1 & 2 & -1 \\ 2 & 0 & -1 & -5 \\ -3 & 1 & 0 & 10 \end{array}\right) &\rightarrow \left(\begin{array}{ccc|c} 2 & 0 & -1 & -5 \\ 0 & 1 & 2 & -1 \\ -3 & 1 & 0 & 10 \end{array}\right) \\
&\rightarrow \left(\begin{array}{ccc|c} 1 & 0 & -1/2 & -5/2 \\ 0 & 1 & 2 & -1 \\ -3 & 1 & 0 & 10 \end{array}\right) \\
&\rightarrow \left(\begin{array}{ccc|c} 1 & 0 & -1/2 & -5/2 \\ 0 & 1 & 2 & -1 \\ 0 & 1 & -3/2 & 5/2 \end{array}\right) \\
&\rightarrow \left(\begin{array}{ccc|c} 1 & 0 & -1/2 & -5/2 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & -7/2 & 7/2 \end{array}\right) \\
&\rightarrow \left(\begin{array}{ccc|c} 1 & 0 & -1/2 & -5/2 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 1 & -1 \end{array}\right) \\
&\rightarrow \left(\begin{array}{ccc|c} 1 & 0 & -1/2 & -5/2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{array}\right) \\
&\rightarrow \left(\begin{array}{ccc|c} 1 & 0 & 0 & -3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{array}\right).
\end{aligned}$$

Thus

$$p = -3f_1 + f_2 - f_3,$$

and has coordinates

$$(p)_B = \begin{pmatrix} -3 \\ 1 \\ -1 \end{pmatrix}.$$

So

$$\begin{aligned}
T_I(p) &= I(p)_B \\
&= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -3 \\ 1 \\ -1 \end{pmatrix} \\
&= \begin{pmatrix} -3 \\ 1 \\ -1 \end{pmatrix}.
\end{aligned}$$

Thus the vector

$$\begin{pmatrix} -3 \\ 1 \\ -1 \end{pmatrix} \in \mathbb{R}^3$$

is the image of p under T_I, that is

$$-x^2 - 5x + 10 \xrightarrow{T_I} \begin{pmatrix} -3 \\ 1 \\ -1 \end{pmatrix}.$$

2. To show that T_I is a linear transformation, we proceed as usual: we must show that T_I interacts nicely with both vector addition and scalar multiplication.

   (a) Given p, $q \in \mathcal{P}_2(\mathbb{R})$ with

   $$\begin{aligned} p(x) &= \alpha_1 f_1(x) + \alpha_2 f_2(x) + \alpha_3 f_3(x) \\ q(x) &= \beta_1 f_1(x) + \beta_2 f_2(x) + \beta_3 f_3(x), \end{aligned}$$

   so that vector $p + q$ is given by

   $$(p+q)(x) = (\alpha_1 + \beta_1) f_1(x) + (\alpha_2 + \beta_2) f_2(x) + (\alpha_3 + \beta_3) f_3(x),$$

   we have the following unique coordinates:

   $$(p)_B = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}, \ (q)_B = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix},$$

   and

   $$(p+q)_B = \begin{pmatrix} \alpha_1 + \beta_1 \\ \alpha_2 + \beta_2 \\ \alpha_3 + \beta_3 \end{pmatrix} = (p)_B + (q)_B.$$

   Thus

   $$\begin{aligned} T_I(p+q) &= I(p+q)_B \\ &= I((p)_B + (q)_B) \\ &= I(p)_B + I(q)_B \\ &= T_I(p) + T_I(q). \end{aligned}$$

   (b) Next, we investigate the action of T_I on λp, $\lambda \in \mathbb{R}$, whose coordinates are given by

   $$(\lambda p)_B = \begin{pmatrix} \lambda\alpha_1 \\ \lambda\alpha_2 \\ \lambda\alpha_3 \end{pmatrix} = \lambda \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \lambda(p)_B.$$

We have

$$\begin{aligned} T_I(\lambda p) &= I(\lambda p)_B \\ &= I(\lambda (p)_B) \\ &= \lambda I(p)_B \\ &= \lambda T_I(p). \end{aligned}$$

Thus T_I interacts nicely with both vector addition and scalar multiplication, and is a linear transformation from $\mathcal{P}_2(\mathbb{R})$ to $\mathbb{R}^3$.

With the previous example in mind, we are now ready to discuss a theorem on the interconnections among linear transformations and matrices:

Theorem. Let V and W be finite dimensional vector spaces with $\dim(V) = n$ and $\dim(W) = m$, and let

$$B = (v_1,\ v_2,\ \ldots,\ v_n)$$

and

$$C = (w_1,\ w_2,\ \ldots,\ w_m)$$

be bases for V and W respectively. Let $T : V \to W$ be *any* linear transformation from V to W. Then there is a *unique* $m \times n$ matrix $A = A_{(B,C)}$ so that, for every $v \in V$,

$$A(v)_B = (T(v))_C,$$

that is, A times the coordinates of V with respect to B yields the coordinates of $T(v)$ with respect to C.

Remark 1. The theorem above gives us an alternate way to think about linear transformations between finite dimensional vector spaces: abusing terminology a bit, *linear transformations are matrices*. More precisely, every matrix can be viewed as a linear transformation between finite dimensional vector spaces, and every linear transformation has realization as a matrix (once a basis has been chosen). Thus the theorem allows us to ignore the distraction of the "decorations" in a particular space, and represent all the interactions algebraically using matrix arithmetic. That is, we can always find a matrix A to do the job of T, expressing all of the calculations in coordinates.

Remark 2. The notation $A_{(B,C)}$, which refers to the matrix realization of linear transformation T, merely indicates that A represents the action of T on V, and converts coordinates with respect to basis B into coordinates with respect to basis C.

Proof. We proceed via construction: given a linear transformation $T : V \to W$ and bases B and C for V, W respectively, we will construct the matrix realization $A = A_{(B,C)}$ for T.

Let $(T(v_i))_C$ be the coordinates of vector $T(v_i) \in W$ with respect to basis C; notice that we may think of $(T(v_i))_C$ as an $m \times 1$ (column) matrix. Set

$$A = \begin{pmatrix} | & | & & | \\ (T(v_1))_C & (T(v_2))_C & \dots & (T(v_n))_C \\ | & | & & | \end{pmatrix},$$

that is A is the $m \times n$ matrix whose ith column is $(T(v_i))_C$.

Now I claim that, for any $v \in V$,

$$(T(v))_C = A(v)_B.$$

To prove the claim, let $v \in V$,

$$v = \alpha_1 v_1 + \ldots + \alpha_n v_n,$$

so that v has coordinates

$$(v)_B = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}.$$

Now

$$\begin{aligned} T(v) &= T(\alpha_1 v_1 + \ldots + \alpha_n v_n) \\ &= \alpha_1 T(v_1) + \ldots + \alpha_n T(v_n), \end{aligned}$$

so that

$$(T(v))_C = \alpha_1 (T(v_1))_C + \ldots + \alpha_n (T(v_n))_C.$$

On the other hand, we see that

$$\begin{aligned} A(v)_B &= \begin{pmatrix} | & | & & | \\ (T(v_1))_C & (T(v_2))_C & \dots & (T(v_n))_C \\ | & | & & | \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \\ &= \alpha_1 (T(v_1))_C + \ldots + \alpha_n (T(v_n))_C. \end{aligned}$$

Thus $A(v)_B = (T(v))_C$, as desired.

To show that A is unique, let A' be any (necessarily $m \times n$) matrix so that $A'(v)_B = (T(v))_C$. For each $v_i \in B$,

$$A'(v_i)_B = (T(v_i))_C.$$

Now v_i is a basis vector, and must have coordinates

$$(v_i)_B = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = e_i,$$

where 1 occurs in the ith position. With this in mind and the fact that $A'(v_i)_B = (T(v_i))_C$, we may calculate the ith column of A': indeed, the only entries of A' that meet a nonzero entry of $(v)_B = e_i$ are those in the ith column, so that

$$A'(v_i)_B = A'e_i = \text{ column i of } A';$$

but since

$$A'(v_i)_B = (T(v_i))_C,$$

the ith column of A' *must be* $(T(v_i))_C$, so that $A = A'$.

Remark. Notice that the proof of the theorem tells us *how to build* A, given transformation T and bases B and C. Indeed, we need merely inspect the coordinates for basis elements of V under the transformation (as indicated by Theorem 3.5); the coordinates for the basis elements of V with respect to basis C of W are the columns of A. In particular,

$$A_{(B,C)} = \begin{pmatrix} | & | & & | \\ (T(v_1))_C & (T(v_2))_C & \dots & (T(v_n))_C \\ | & | & & | \end{pmatrix}.$$

Example. Recall that the vectors $f_1(x) = 2x - 3$, $f_2(x) = x^2 + 1$, and $f_3(x) = 2x^2 - x$ are linearly independent and span the vector space $\mathcal{P}_2(\mathbb{R})$ of all real-valued polynomials of degree no more than 2; of course, $p_1(x) = x^2$, $p_2(x) = x$, and $p_3(x) = 1$ have the same properties. Thus the lists

$$B = (f_1,\ f_2,\ f_3)$$

and

$$\hat{B} = (p_1,\ p_2,\ p_3)$$

are both bases for $\mathcal{P}_2(\mathbb{R})$.

Consider the linear transformation $T : \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}^3$ given by

$$T(\alpha x^2 + \beta x + \delta) := \begin{pmatrix} \alpha \\ \beta \\ \delta \end{pmatrix}.$$

1. Given $p(x) = -x^2 - 5x + 10$, find $T(p)$ and $(T(p))_C$.
2. Find the matrix $A = A_{(B,C)}$ for T with respect to basis B and the standard basis for $\mathbb{R}^3$ (denoted by C).
3. With $p(x) = -x^2 - 5x + 10$, show that

$$A(p)_B = (T(p))_C.$$

4. Find the matrix $\hat{A} = A_{(\hat{B},C)}$ for T with respect to basis $\hat{B}$ and the standard basis for $\mathbb{R}^3$.

5. Show that
$$\hat{A}(p)_B = (T(p))_C.$$

1. Using the definition of T, we see that
$$T(-5x^2 - x + 10) = \begin{pmatrix} -5 \\ -1 \\ 10 \end{pmatrix};$$
since
$$\begin{pmatrix} -5 \\ -1 \\ 10 \end{pmatrix} = -5 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + 10 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$
the coordinates $(T(p))_C$ of $T(p)$ with respect to the standard basis C on $\mathbb{R}^3$ are given by
$$T(p)_C = \begin{pmatrix} -5 \\ -1 \\ 10 \end{pmatrix}.$$

2. To find matrix A, we only need to find the coordinates of $T(f_i)$ with respect to basis C. Thus

we evaluate T at the basis vectors from B:

$$\begin{aligned}
T(f_1) &= T(2x-3) \\
&= \begin{pmatrix} 0 \\ 2 \\ -3 \end{pmatrix} \\
&= 0\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 2\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - 3\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}; \\
T(f_2) &= T(x^2+1) \\
&= \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \\
&= 1\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 0\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + 1\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}; \\
T(f_3) &= T(2x^2-x) \\
&= \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} \\
&= 2\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + 0\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.
\end{aligned}$$

Thus we have the following coordinates for the images of the basis vectors from B:

$$\begin{aligned}
(T(f_1))_C &= \begin{pmatrix} 0 \\ 2 \\ -3 \end{pmatrix}; \\
(T(f_2))_C &= \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}; \\
(T(f_3))_C &= \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}.
\end{aligned}$$

The matrix representing T with respect to bases B and C is

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

3. To find the image of $p(x) = -5x^2 - x + 10$ under the action of T using the matrix A, we recall that A is the matrix so that

$$A(p)_B = (T(p))_C;$$

that is, A times the coordinate vector of p with respect to B will yield the coordinates (with respect to basis C) of $T(p)$.

As we saw earlier, the coordinates of $p(x)$ with respect to basis B are given by

$$(p)_B = \begin{pmatrix} -3 \\ 1 \\ -1 \end{pmatrix}.$$

Thus we caclulate

$$\begin{aligned} A(p)_B &= \begin{pmatrix} 0 & 1 & 2 \\ 2 & 0 & -1 \\ -3 & 1 & 0 \end{pmatrix} \begin{pmatrix} -3 \\ 1 \\ -1 \end{pmatrix} \\ &= \begin{pmatrix} -1 \\ -5 \\ 10 \end{pmatrix}. \end{aligned}$$

Notice that this matches up precisely with the vector $(T(p))_C$ that we found in part 1.

4. To find $\hat{A}$, we need to consider the action of T on the standard basis for $\mathcal{P}_2(\mathbb{R})$. Accordingly, we have

$$\begin{aligned} T(p_1) &= T(x^2) \\ &= \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}; \\ T(p_2) &= T(x) \\ &= \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}; \\ T(p_3) &= T(1) \\ &= \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \end{aligned}$$

Thus $\hat{A}$ is given by

$$\hat{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

5. Vector $p(x)$ has coordinates with respect to basis $\hat{B}$ given by

$$(p)_{\hat{B}} = \begin{pmatrix} -1 \\ -5 \\ 10 \end{pmatrix}.$$

Thus

$$\begin{aligned} \hat{A}(p)_{\hat{B}} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 \\ -5 \\ 10 \end{pmatrix} \\ &= \begin{pmatrix} -1 \\ -5 \\ 10 \end{pmatrix}, \end{aligned}$$

which again matches up precisely with $(T(p))_C$.

Example. Recall that the vector space $\mathfrak{sl}(2,\mathbb{R})$ of 2×2 trace 0 matrices has basis

$$B = (e_1,\ e_2,\ e_3),$$

where

$$e_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},\ e_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},\ \text{and } e_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Let $T : \mathfrak{sl}(2,\mathbb{R}) \to \mathfrak{sl}(2,\mathbb{R})$ be the linear transformation with matrix (with respect to B)

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

Describe the action of T on any vector

$$v = \begin{pmatrix} \alpha & \beta \\ \gamma & -\alpha \end{pmatrix} \in \mathfrak{sl}(2,\mathbb{R}).$$

Recall that the action of T on $v \in \mathfrak{sl}(2,\mathbb{R})$ is completely determined by its action on the basis B. Thus we simply need to describe the action of T on each of e_1, e_2, and e_3.

We know that

$$\begin{aligned}(T(e_1))_B &= A(e_1)_B \\ &= \begin{pmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 1 & 0 & 0\end{pmatrix}\begin{pmatrix}1\\0\\0\end{pmatrix} \\ &= \begin{pmatrix}0\\0\\1\end{pmatrix} \\ &= (e_3)_B;\end{aligned}$$

$$\begin{aligned}(T(e_2))_B &= A(e_2)_B \\ &= \begin{pmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 1 & 0 & 0\end{pmatrix}\begin{pmatrix}0\\1\\0\end{pmatrix} \\ &= \begin{pmatrix}1\\0\\0\end{pmatrix} \\ &= (e_1)_B;\end{aligned}$$

and

$$\begin{aligned}(T(e_3))_B &= A(e_3)_B \\ &= \begin{pmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 1 & 0 & 0\end{pmatrix}\begin{pmatrix}0\\0\\1\end{pmatrix} \\ &= \begin{pmatrix}0\\1\\0\end{pmatrix} \\ &= (e_2)_B.\end{aligned}$$

Reverting from coordinate representation to $\mathfrak{sl}(2,\mathbb{R})$ itself, we see that

$$\begin{aligned}T\left(\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\right) &= \begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}; \\ T\left(\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}\right) &= \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}; \text{ and} \\ T\left(\begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}\right) &= \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}.\end{aligned}$$

Thus the action of T on any vector

$$v = \begin{pmatrix} \alpha & \beta \\ \gamma & -\alpha \end{pmatrix}$$

is given by

$$\begin{aligned} T(v) &= T\left(\begin{pmatrix} \alpha & \beta \\ \gamma & -\alpha \end{pmatrix} \right) \\ &= T\left(\alpha \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \beta \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + \gamma \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \right) \\ &= \alpha T\left(\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right) + \beta T\left(\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \right) + \gamma T\left(\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \right) \\ &= \alpha \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + \beta \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \gamma \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} \beta & \gamma \\ \alpha & -\beta \end{pmatrix}. \end{aligned}$$

As indicated by the last example, linear transformations from a space V to itself are particularly important; below, we give such transformations a name.

Definition 3.67. A linear transformation from a vector space V to itself is called a *linear operator*.

Note that the linear transformation T above,

$$T\left(\begin{pmatrix} \alpha & \beta \\ \gamma & -\alpha \end{pmatrix} \right) = \begin{pmatrix} \beta & \gamma \\ \alpha & -\beta \end{pmatrix},$$

is an example of a linear operator, since $T : \mathfrak{sl}(2, \mathbb{R}) \to \mathfrak{sl}(2, \mathbb{R})$.

Change of Basis

Let V be a vector space with bases

$$B = (v_1,\ v_2,\ \ldots,\ v_n)$$

and

$$C = (w_1,\ w_2,\ \ldots,\ w_n).$$

We may happen to know the coordinates for $v \in V$ with respect to basis B, say

$$(v)_B = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}.$$

Of course, v has coordinates with respect to C as well, say

$$(v)_C = \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_n \end{pmatrix}.$$

We would like to have a process for finding $(v)_C$ if we already know $(v)_B$, that is we would like to be able to convert coordinates from one basis to another. This process of converting the coordinates of a vector from one basis to another can be accomplished rapidly using matrix multiplication.

Indeed, we can construct a matrix $A = A_{(B,C)}$ for the identity transformation $T_I : V \to V$, $T_I(v) = v$. In particular, since $T_I(v_i) = v_i$, the matrix A for T_I has form

$$A = \begin{pmatrix} | & | & & | \\ (v_1)_C & (v_2)_C & \dots & (v_n)_C \\ | & | & & | \end{pmatrix}.$$

Using the theorem above, we see that, for any $v \in V$,

$$A(v)_B = (T_I(v))_C = (v)_C,$$

that is, matrix A *converts coordinates for* v *with respect to basis* B *into coordinates with respect to basis* C.

Definition. Let V be an n dimensional vector space with bases

$$B = (v_1,\ v_2,\ \dots,\ v_n)$$

and

$$C = (w_1,\ w_2,\ \dots,\ w_n).$$

The $n \times n$ matrix $A = A_{(B,C)}$ representing the identity transformation T_I is called the *transition matrix* from coordinates with respect to B to coordinates with respect to C. The ith column of A consists of the coordinates of v_i with respect to to basis C, that is

$$A = \begin{pmatrix} | & | & & | \\ (v_1)_C & (v_2)_C & \dots & (v_n)_C \\ | & | & & | \end{pmatrix}.$$

The following theorem reiterates the important feature of a transition matrix which we discovered above:

Theorem. Let B and C be bases for a vector space V. Then $A = A_{(B,C)}$ is the transition matrix from coordinates in terms of B to coordinates in terms of C if and only if

$$A(v)_B = (v)_C$$

for all $v \in V$.

Remark. A transition matrix is the matrix for a particular linear operator–the identity transformation T_I. A transition matrix does not change a vector–it simply encodes a "change of basis". Again a transition matrix rewrites coordinates in terms of one basis with respect to the new basis.

Example. Given vectors

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \ e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \ v_1 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \text{ and } v_2 = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

vector space $\mathbb{R}^2$ has bases

$$B = (e_1, \ e_2)$$

and

$$C = (v_1, \ v_2).$$

1. Find the transition matrix A from coordinates with respect to B to coordinates with respect to C.
2. Use the transition matrix to calculate the coordinates of point $p = (-7, 3)$ in $\mathbb{R}^2$ with respect to basis C.

1. To find the 2×2 matrix A that rewrites coordinates in terms of basis B as coordinates in terms of basis C, we must write vectors in B as linear combinations of vectors in C. It is easy to see that

$$\begin{aligned} v_1 &= 3e_1 + e_2 \\ v_2 &= 2e_1 + 2e_2. \end{aligned}$$

Solving for e_1 and e_2 in the first pair of equations, we have

$$\begin{aligned} e_1 &= \tfrac{1}{2}\, v_1 - \tfrac{1}{4}\, v_2 \\ e_2 &= -\tfrac{1}{2}\, v_1 + \tfrac{3}{4}\, v_2. \end{aligned}$$

In particular, this gives us the coordinates of e_1 and e_2 with respect to the new basis:

$$(e_1)_C = \begin{pmatrix} \dfrac{1}{2} \\ \\ -\dfrac{1}{4} \end{pmatrix} \text{ and } (e_2)_C = \begin{pmatrix} -\dfrac{1}{2} \\ \\ \dfrac{3}{4} \end{pmatrix}.$$

Thus the transition matrix is given by

$$A = \begin{pmatrix} | & | \\ (e_1)_C & (e_2)_C \\ | & | \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{4} & \frac{3}{4} \end{pmatrix}.$$

2. Since A is a transition matrix, we know that

$$A(p)_B = (p)_C.$$

Since the coordinates for $p = (-7, 3)$ with respect to basis B are just

$$(p)_B = \begin{pmatrix} -7 \\ 3 \end{pmatrix},$$

the coordinates of p with respect to basis C are

$$\begin{aligned} (p)_C &= A(p)_B \\ &= \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{4} & \frac{3}{4} \end{pmatrix} \begin{pmatrix} -7 \\ 3 \end{pmatrix} \\ &= \begin{pmatrix} -5 \\ 4 \end{pmatrix}. \end{aligned}$$

The calculation above is easy to verify–indeed, you should check that $p = -5v_1 + 4v_2$.

Example. Given the vector space $\mathfrak{sl}(2, \mathbb{R})$ of 2×2 trace 0 matrices, basis

$$B = (e_1,\ e_2,\ e_3),$$

we may think of matrix

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

as a transition matrix from coordinates in terms of B to coordinates in terms of a new basis C. Find C.

Column i of A consists of the coordinates of e_i with respect to basis $C = (w_1,\ w_2,\ w_3)$. Since the first column of A is

$$\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

we see that

$$e_1 = 0w_1 + 0w_2 + w_3 = w_3.$$

Similarly,

$$e_2 = w_1 + 0w_2 + 0w_3 = w_1,$$

and

$$e_3 = 0w_1 + w_2 + 0w_3.$$

Thus basis C is just

$$C = (e_2,\ e_3,\ e_1).$$

Let's check that the calculations above make sense, with a particular matrix in $\mathfrak{sl}(2, \mathbb{R})$: given

$$v = \begin{pmatrix} -1 & 2 \\ 4 & 1 \end{pmatrix},$$

it is easy to see that v has coordinates

$$(v)_B = \begin{pmatrix} -1 \\ 2 \\ 4 \end{pmatrix}$$

and

$$(v)_C = \begin{pmatrix} 2 \\ 4 \\ -1 \end{pmatrix}.$$

Now applying A to $(v)_B$, we have

$$\begin{aligned} A(v)_B &= \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} -1 \\ 2 \\ 4 \end{pmatrix} \\ &= \begin{pmatrix} 2 \\ 4 \\ -1 \end{pmatrix} \\ &= (v)_C. \end{aligned}$$

Thus A simply converts coordinates for v in terms of B to coordinates in terms of C.

In addition to finding the transition matrix $A = A_{(B,C)}$, we can also create the transition matrix $A' = A_{(C,B)}$ that converts coordinates in terms of C to coordinates in terms of B. A' reverses the action of A–that is,

$$A'A(v)_B = A'(v)_C = (v)_B.$$

Thus the following theorem is not too surprising:

Theorem. The transition matrix $A = A_{(B,C)}$ is invertible, and A^{-1} is given by $A^{-1} = A_{(C,B)}$.

We can use this theorem to our advantage if we wish to find a nicer "version" of the matrix for an operator T. In particular, suppose that $A = A_{(B,B)}$ is the matrix for T with respect to basis B, and let X be the transition matrix from coordinates in terms of B to coordinates in terms of C, so that X^{-1} converts coordinates from C to B. Then I claim that XAX^{-1} is the matrix for T with respect to C, that is

$$A_{(C,C)} = XAX^{-1}.$$

To check the claim, we need to verify that

$$XAX^{-1}(v)_C = (T(v_C));$$

since X^{-1} converts coordinates from C to B, we have

$$\begin{aligned} XAX^{-1}(v)_C &= XA(X^{-1}(v)_C) \\ &= X(A(v)_B) \\ &= X(T(v))_B \\ &= (T(v))_C \end{aligned}$$

as desired, since X converts coordinates from B to C.

Example. Recall that the vector space $\mathbb{R}^2$ has bases

$$B = (e_1,\ e_2)$$

and

$$C = (v_1,\ v_2).$$

Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the operator with matrix

$$A = A_{(C,C)} = \begin{pmatrix} 5/2 & 1 \\ -3/4 & 1/2 \end{pmatrix}.$$

Find the matrix of T with respect to B and describe the action of T on $x \in \mathbb{R}^2$.

We simply need to know the transition matrices X from C to B, and its inverse, the transition matrix from B to C. Of course, we already know that

$$X = \begin{pmatrix} 1/2 & -1/2 \\ 1/4 & 3/4 \end{pmatrix} \text{ and } X^{-1} = \begin{pmatrix} 3 & 2 \\ 1 & 2 \end{pmatrix}.$$

Thus

$$\begin{aligned} A_{(B,B)} &= XAX^{-1} \\ &= \begin{pmatrix} 1/2 & -1/2 \\ 1/4 & 3/4 \end{pmatrix} \begin{pmatrix} 5/2 & 1 \\ -3/4 & 1/2 \end{pmatrix} \begin{pmatrix} 3 & 2 \\ 1 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

Now it is much easier to see how T treats vectors in $\mathbb{R}^2$: indeed, since $x = (x)_B$ in this example, we see that

$$\begin{aligned} T(x) &= (T(x))_B \\ &= A_{(B,B)}x \\ &= \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= \begin{pmatrix} 2x_1 \\ x_2 \end{pmatrix}. \end{aligned}$$

Matrix for the Composition of Linear Transformations

Recall that, if $T \in \mathcal{L}(V, U)$ and $S \in \mathcal{L}(U, W)$, then the *product* ST defined by

$$ST(v) = S(T(v))$$

is a linear transformation, and

$$ST : V \to W,$$

that is $ST \in \mathcal{L}(V, W)$. We think of ST as a composition of the transformations T and S.

Now choosing basis B for V, C for U, and D for W, we may find the matrix representations for T and S, say $A_{(B,C)}$ for T and $A_{(C,D)}$ for S. We naturally ask:

> What is the relationship between the matrix representation for the composition ST and the matrix representations for T and S?

As one might suspect, given our knowledge of matrix arithmetic and the interconnections with linear transformations, the matrix for a composition is the *product* of the matrices for the component transformations:

Theorem. Let V, U and W be finite dimensional vector spaces with bases B, C, and D respectively. Let $T : V \to U$ and $S : U \to W$ be linear transformations with corresponding matrix representations $A_{(B,C)}$ and $A_{(C,D)}$ respectively. Then the matrix $A_{(B,D)}$ for the linear transformation $ST : V \to W$ is the product

$$A_{(B,D)} = A_{(C,D)} A_{(B,C)}.$$

Remark. We have already seen an example of the use of the theorem above. In Unit 3, Section 1, we considered the linear transformation $T_M : \mathbb{R}^2 \to \mathbb{R}^3$ with associated matrix

$$M = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}$$

and $T_N : \mathbb{R}^3 \to \mathbb{R}^1$ with associated matrix

$$N = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix};$$

then the transformation

$$T_N T_M : \mathbb{R}^2 \to \mathbb{R}^1$$

has matrix

$$NM = \begin{pmatrix} 2 & 2 \end{pmatrix}.$$