LINEAR ALGEBRA

W W L CHEN

Chapter 8

LINEAR TRANSFORMATIONS

8.1. Euclidean Linear Transformations

By a transformation from $\mathbb{R}^n$ into $\mathbb{R}^m$, we mean a function of the type $T : \mathbb{R}^n \to \mathbb{R}^m$, with domain $\mathbb{R}^n$ and codomain $\mathbb{R}^m$. For every vector $\mathbf{x} \in \mathbb{R}^n$, the vector $T(\mathbf{x}) \in \mathbb{R}^m$ is called the image of $\mathbf{x}$ under the transformation T, and the set

$$R(T) = \{T(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^n\},$$

of all images under T, is called the range of the transformation T.

Remark. For our convenience later, we have chosen to use $R(T)$ instead of the usual $T(\mathbb{R}^n)$ to denote the range of the transformation T.

For every $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$, we can write

$$T(\mathbf{x}) = T(x_1, \ldots, x_n) = (y_1, \ldots, y_m).$$

Here, for every $i = 1, \ldots, m$, we have

$$y_i = T_i(x_1, \ldots, x_n), \tag{1}$$

where $T_i : \mathbb{R}^n \to \mathbb{R}$ is a real valued function.

Definition. A transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is called a linear transformation if there exists a real matrix

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix}$$

such that for every $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$, we have $T(x_1, \ldots, x_n) = (y_1, \ldots, y_m)$, where

$$\begin{aligned} y_1 &= a_{11}x_1 + \ldots + a_{1n}x_n, \\ &\vdots \\ y_m &= a_{m1}x_1 + \ldots + a_{mn}x_n, \end{aligned}$$

or, in matrix notation,

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}. \tag{2}$$

The matrix A is called the standard matrix for the linear transformation T.

Remarks. (1) In other words, a transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is linear if the equation (1) for every $i = 1, \ldots, m$ is linear.

(2) If we write $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{y} \in \mathbb{R}^m$ as column matrices, then (2) can be written in the form $\mathbf{y} = A\mathbf{x}$, and so the linear transformation T can be interpreted as multiplication of $\mathbf{x} \in \mathbb{R}^n$ by the standard matrix A.

Definition. A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is said to be a linear operator if $n = m$. In this case, we say that T is a linear operator on $\mathbb{R}^n$.

Example 8.1.1. The linear transformation $T : \mathbb{R}^5 \to \mathbb{R}^3$, defined by the equations

$$\begin{aligned} y_1 &= 2x_1 + 3x_2 + 5x_3 + 7x_4 - 9x_5, \\ y_2 &= \phantom{2x_1 + {}} 3x_2 + 4x_3 \phantom{{}+ 7x_4} + 2x_5, \\ y_3 &= x_1 \phantom{{}+ 3x_2} + 3x_3 - 2x_4 \phantom{{}- 9x_5}, \end{aligned}$$

can be expressed in matrix form as

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 2 & 3 & 5 & 7 & -9 \\ 0 & 3 & 4 & 0 & 2 \\ 1 & 0 & 3 & -2 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix}.$$

If $(x_1, x_2, x_3, x_4, x_5) = (1, 0, 1, 0, 1)$, then

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 2 & 3 & 5 & 7 & -9 \\ 0 & 3 & 4 & 0 & 2 \\ 1 & 0 & 3 & -2 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ 6 \\ 4 \end{pmatrix},$$

so that $T(1, 0, 1, 0, 1) = (-2, 6, 4)$.

Example 8.1.2. Suppose that A is the zero $m \times n$ matrix. The linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$, where $T(\mathbf{x}) = A\mathbf{x}$ for every $\mathbf{x} \in \mathbb{R}^n$, is the zero transformation from $\mathbb{R}^n$ into $\mathbb{R}^m$. Clearly $T(\mathbf{x}) = \mathbf{0}$ for every $\mathbf{x} \in \mathbb{R}^n$.

Example 8.1.3. Suppose that I is the identity $n \times n$ matrix. The linear operator $T : \mathbb{R}^n \to \mathbb{R}^n$, where $T(\mathbf{x}) = I\mathbf{x}$ for every $\mathbf{x} \in \mathbb{R}^n$, is the identity operator on $\mathbb{R}^n$. Clearly $T(\mathbf{x}) = \mathbf{x}$ for every $\mathbf{x} \in \mathbb{R}^n$.

PROPOSITION 8A. *Suppose that $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, and that $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is the standard basis for $\mathbb{R}^n$. Then the standard matrix for T is given by*

$$A = (\,T(\mathbf{e}_1) \quad \ldots \quad T(\mathbf{e}_n)\,),$$

where $T(\mathbf{e}_j)$ is a column matrix for every $j = 1, \ldots, n$.

PROOF. This follows immediately from (2). ◯

8.2. Linear Operators on $\mathbb{R}^2$

In this section, we consider the special case when $n = m = 2$, and study linear operators on $\mathbb{R}^2$. For every $\mathbf{x} \in \mathbb{R}^2$, we shall write $\mathbf{x} = (x_1, x_2)$.

EXAMPLE 8.2.1. Consider reflection across the x_2-axis, so that $T(x_1, x_2) = (-x_1, x_2)$. Clearly we have

$$T(\mathbf{e}_1) = \begin{pmatrix} -1 \\ 0 \end{pmatrix} \quad \text{and} \quad T(\mathbf{e}_2) = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

and so it follows from Proposition 8A that the standard matrix is given by

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

It is not difficult to see that the standard matrices for reflection across the x_1-axis and across the line $x_1 = x_2$ are given respectively by

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{and} \quad A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Also, the standard matrix for reflection across the origin is given by

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

We give a summary in the table below:

Linear operator	Equations	Standard matrix
Reflection across x_2-axis	$\begin{cases} y_1 = -x_1 \\ y_2 = x_2 \end{cases}$	$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$
Reflection across x_1-axis	$\begin{cases} y_1 = x_1 \\ y_2 = -x_2 \end{cases}$	$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$
Reflection across $x_1 = x_2$	$\begin{cases} y_1 = x_2 \\ y_2 = x_1 \end{cases}$	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$
Reflection across origin	$\begin{cases} y_1 = -x_1 \\ y_2 = -x_2 \end{cases}$	$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$

EXAMPLE 8.2.2. For orthogonal projection onto the x_1-axis, we have $T(x_1, x_2) = (x_1, 0)$, with standard matrix

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

Similarly, the standard matrix for orthogonal projection onto the x_2-axis is given by

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

We give a summary in the table below:

Linear operator	Equations	Standard matrix
Orthogonal projection onto x_1-axis	$\begin{cases} y_1 = x_1 \\ y_2 = 0 \end{cases}$	$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$
Orthogonal projection onto x_2-axis	$\begin{cases} y_1 = 0 \\ y_2 = x_2 \end{cases}$	$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$

EXAMPLE 8.2.3. For anticlockwise rotation by an angle θ, we have $T(x_1, x_2) = (y_1, y_2)$, where

$$y_1 + \mathrm{i}y_2 = (x_1 + \mathrm{i}x_2)(\cos\theta + \mathrm{i}\sin\theta),$$

and so

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

It follows that the standard matrix is given by

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

We give a summary in the table below:

Linear operator	Equations	Standard matrix
Anticlockwise rotation by angle θ	$\begin{cases} y_1 = x_1\cos\theta - x_2\sin\theta \\ y_2 = x_1\sin\theta + x_2\cos\theta \end{cases}$	$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$

EXAMPLE 8.2.4. For contraction or dilation by a non-negative scalar k, we have $T(x_1, x_2) = (kx_1, kx_2)$, with standard matrix

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

The operator is called a contraction if $0 < k < 1$ and a dilation if $k > 1$, and can be extended to negative values of k by noting that for $k < 0$, we have

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

This describes contraction or dilation by non-negative scalar $-k$ followed by reflection across the origin. We give a summary in the table below:

Linear operator	Equations	Standard matrix
Contraction or dilation by factor k	$\begin{cases} y_1 = kx_1 \\ y_2 = kx_2 \end{cases}$	$\begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}$

EXAMPLE 8.2.5. For expansion or compression in the x_1-direction by a positive factor k, we have $T(x_1, x_2) = (kx_1, x_2)$, with standard matrix

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

This can be extended to negative values of k by noting that for $k < 0$, we have

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

This describes expansion or compression in the x_1-direction by positive factor $-k$ followed by reflection across the x_2-axis. Similarly, for expansion or compression in the x_2-direction by a non-zero factor k, we have the standard matrix

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

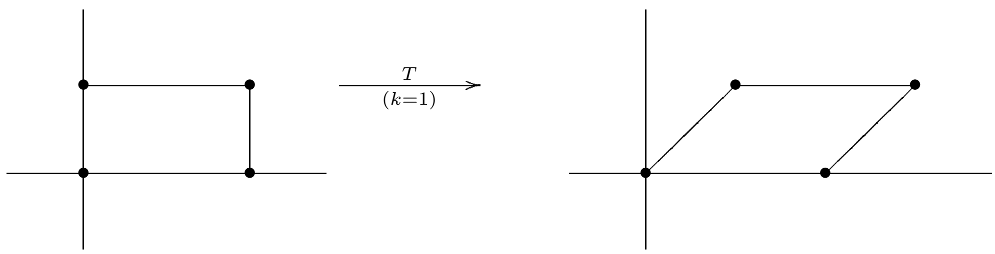
We give a summary in the table below:

Linear operator	Equations	Standard matrix
Expansion or compression in x_1-direction	$\begin{cases} y_1 = kx_1 \\ y_2 = x_2 \end{cases}$	$\begin{pmatrix} k & 0 \\ 0 & 1 \end{pmatrix}$
Expansion or compression in x_2-direction	$\begin{cases} y_1 = x_1 \\ y_2 = kx_2 \end{cases}$	$\begin{pmatrix} 1 & 0 \\ 0 & k \end{pmatrix}$

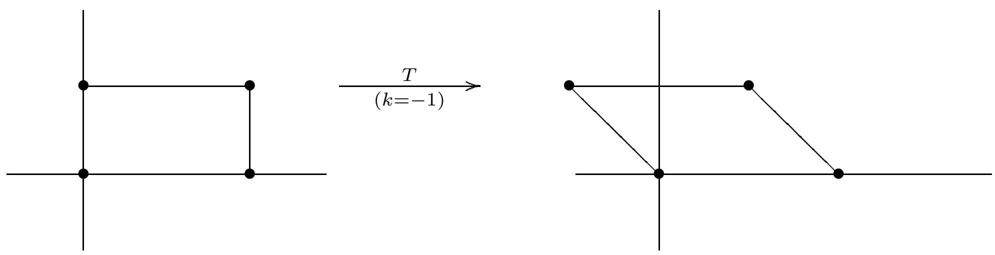
EXAMPLE 8.2.6. For shears in the x_1-direction with factor k, we have $T(x_1, x_2) = (x_1 + kx_2, x_2)$, with standard matrix

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

For the case $k = 1$, we have the following.



For the case $k = -1$, we have the following.



Similarly, for shears in the x_2-direction with factor k, we have standard matrix

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

We give a summary in the table below:

Linear operator	Equations	Standard matrix
Shear in x_1-direction	$\begin{cases} y_1 = x_1 + kx_2 \\ y_2 = x_2 \end{cases}$	$\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$
Shear in x_2-direction	$\begin{cases} y_1 = x_1 \\ y_2 = kx_1 + x_2 \end{cases}$	$\begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix}$

EXAMPLE 8.2.7. Consider a linear operator $T : \mathbb{R}^2 \to \mathbb{R}^2$ which consists of a reflection across the x_2-axis, followed by a shear in the x_1-direction with factor 3 and then reflection across the x_1-axis. To find the standard matrix, consider the effect of T on a standard basis $\{\mathbf{e}_1, \mathbf{e}_2\}$ of $\mathbb{R}^2$. Note that

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} -1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} -1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} -1 \\ 0 \end{pmatrix} = T(\mathbf{e}_1),$$
$$\mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} 3 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} 3 \\ -1 \end{pmatrix} = T(\mathbf{e}_2),$$

so it follows from Proposition 8A that the standard matrix for T is

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

Let us summarize the above and consider a few special cases. We have the following table of invertible linear operators with $k \neq 0$. Clearly, if A is the standard matrix for an invertible linear operator T, then the inverse matrix A^{-1} is the standard matrix for the inverse linear operator T^{-1}.

Linear operator T	Standard matrix A	Inverse matrix A^{-1}	Linear operator T^{-1}
Reflection across line $x_1=x_2$	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$	Reflection across line $x_1=x_2$
Expansion or compression in x_1-direction	$\begin{pmatrix} k & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} k^{-1} & 0 \\ 0 & 1 \end{pmatrix}$	Expansion or compression in x_1-direction
Expansion or compression in x_2-direction	$\begin{pmatrix} 1 & 0 \\ 0 & k \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & k^{-1} \end{pmatrix}$	Expansion or compression in x_2-direction
Shear in x_1-direction	$\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & -k \\ 0 & 1 \end{pmatrix}$	Shear in x_1-direction
Shear in x_2-direction	$\begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ -k & 1 \end{pmatrix}$	Shear in x_2-direction

Next, let us consider the question of elementary row operations on 2×2 matrices. It is not difficult to see that an elementary row operation performed on a 2×2 matrix A has the effect of multiplying the matrix A by some elementary matrix E to give the product EA. We have the following table.

Elementary row operation	Elementary matrix E
Interchanging the two rows	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$
Multiplying row 1 by non-zero factor k	$\begin{pmatrix} k & 0 \\ 0 & 1 \end{pmatrix}$
Multiplying row 2 by non-zero factor k	$\begin{pmatrix} 1 & 0 \\ 0 & k \end{pmatrix}$
Adding k times row 2 to row 1	$\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$
Adding k times row 1 to row 2	$\begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix}$

Now, we know that any invertible matrix A can be reduced to the identity matrix by a finite number of elementary row operations. In other words, there exist a finite number of elementary matrices $E_1, \ldots, E_s$ of the types above with various non-zero values of k such that

$$E_s \ldots E_1 A = I,$$

so that

$$A = E_1^{-1} \ldots E_s^{-1}.$$

We have proved the following result.

PROPOSITION 8B. *Suppose that the linear operator $T : \mathbb{R}^2 \to \mathbb{R}^2$ has standard matrix A, where A is invertible. Then T is the product of a succession of finitely many reflections, expansions, compressions and shears.*

In fact, we can prove the following result concerning images of straight lines.

PROPOSITION 8C. *Suppose that the linear operator $T : \mathbb{R}^2 \to \mathbb{R}^2$ has standard matrix A, where A is invertible. Then*
(a) the image under T of a straight line is a straight line;
(b) the image under T of a straight line through the origin is a straight line through the origin; and
(c) the images under T of parallel straight lines are parallel straight lines.

PROOF. Suppose that $T(x_1, x_2) = (y_1, y_2)$. Since A is invertible, we have $\mathbf{x} = A^{-1}\mathbf{y}$, where

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad \text{and} \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

The equation of a straight line is given by $\alpha x_1 + \beta x_2 = \gamma$ or, in matrix form, by

$$(\,\alpha \quad \beta\,) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (\,\gamma\,).$$

Hence

$$(\,\alpha \quad \beta\,) A^{-1} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = (\,\gamma\,).$$

Let

$$\begin{pmatrix} \alpha' & \beta' \end{pmatrix} = \begin{pmatrix} \alpha & \beta \end{pmatrix} A^{-1}.$$

Then

$$\begin{pmatrix} \alpha' & \beta' \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \gamma \end{pmatrix}.$$

In other words, the image under T of the straight line $\alpha x_1 + \beta x_2 = \gamma$ is $\alpha' y_1 + \beta' y_2 = \gamma$, clearly another straight line. This proves (a). To prove (b), note that straight lines through the origin correspond to $\gamma = 0$. To prove (c), note that parallel straight lines correspond to different values of γ for the same values of α and β. ○

8.3. Elementary Properties of Euclidean Linear Transformations

In this section, we establish a number of simple properties of euclidean linear transformations.

PROPOSITION 8D. *Suppose that $T_1 : \mathbb{R}^n \to \mathbb{R}^m$ and $T_2 : \mathbb{R}^m \to \mathbb{R}^k$ are linear transformations. Then $T = T_2 \circ T_1 : \mathbb{R}^n \to \mathbb{R}^k$ is also a linear transformation.*

PROOF. Since T_1 and T_2 are linear transformations, they have standard matrices A_1 and A_2 respectively. In other words, we have $T_1(\mathbf{x}) = A_1\mathbf{x}$ for every $\mathbf{x} \in \mathbb{R}^n$ and $T_2(\mathbf{y}) = A_2\mathbf{y}$ for every $\mathbf{y} \in \mathbb{R}^m$. It follows that $T(\mathbf{x}) = T_2(T_1(\mathbf{x})) = A_2A_1\mathbf{x}$ for every $\mathbf{x} \in \mathbb{R}^n$, so that T has standard matrix A_2A_1. ○

EXAMPLE 8.3.1. Suppose that $T_1 : \mathbb{R}^2 \to \mathbb{R}^2$ is anticlockwise rotation by $\pi/2$ and $T_2 : \mathbb{R}^2 \to \mathbb{R}^2$ is orthogonal projection onto the x_1-axis. Then the respective standard matrices are

$$A_1 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad A_2 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

It follows that the standard matrices for $T_2 \circ T_1$ and $T_1 \circ T_2$ are respectively

$$A_2A_1 = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad A_1A_2 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Hence $T_2 \circ T_1$ and $T_1 \circ T_2$ are not equal.

EXAMPLE 8.3.2. Suppose that $T_1 : \mathbb{R}^2 \to \mathbb{R}^2$ is anticlockwise rotation by θ and $T_2 : \mathbb{R}^2 \to \mathbb{R}^2$ is anticlockwise rotation by ϕ. Then the respective standard matrices are

$$A_1 = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \quad \text{and} \quad A_2 = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}.$$

It follows that the standard matrix for $T_2 \circ T_1$ is

$$A_2A_1 = \begin{pmatrix} \cos\phi\cos\theta - \sin\phi\sin\theta & -\cos\phi\sin\theta - \sin\phi\cos\theta \\ \sin\phi\cos\theta + \cos\phi\sin\theta & \cos\phi\cos\theta - \sin\phi\sin\theta \end{pmatrix} = \begin{pmatrix} \cos(\phi+\theta) & -\sin(\phi+\theta) \\ \sin(\phi+\theta) & \cos(\phi+\theta) \end{pmatrix}.$$

Hence $T_2 \circ T_1$ is anticlockwise rotation by $\phi + \theta$.

EXAMPLE 8.3.3. The reader should check that in $\mathbb{R}^2$, reflection across the x_1-axis followed by reflection across the x_2-axis gives reflection across the origin.

Linear transformations that map distinct vectors to distinct vectors are of special importance.

Definition. A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is said to be one-to-one if for every $\mathbf{x}', \mathbf{x}'' \in \mathbb{R}^n$, we have $\mathbf{x}' = \mathbf{x}''$ whenever $T(\mathbf{x}') = T(\mathbf{x}'')$.

Example 8.3.4. If we consider linear operators $T : \mathbb{R}^2 \to \mathbb{R}^2$, then T is one-to-one precisely when the standard matrix A is invertible. To see this, suppose first of all that A is invertible. If $T(\mathbf{x}') = T(\mathbf{x}'')$, then $A\mathbf{x}' = A\mathbf{x}''$. Multiplying on the left by A^{-1}, we obtain $\mathbf{x}' = \mathbf{x}''$. Suppose next that A is not invertible. Then there exists $\mathbf{x} \in \mathbb{R}^2$ such that $\mathbf{x} \neq \mathbf{0}$ and $A\mathbf{x} = \mathbf{0}$. On the other hand, we clearly have $A\mathbf{0} = \mathbf{0}$. It follows that $T(\mathbf{x}) = T(\mathbf{0})$, so that T is not one-to-one.

PROPOSITION 8E. *Suppose that the linear operator $T : \mathbb{R}^n \to \mathbb{R}^n$ has standard matrix A. Then the following statements are equivalent:*
(a) The matrix A is invertible.
(b) The linear operator T is one-to-one.
(c) The range of T is $\mathbb{R}^n$; in other words, $R(T) = \mathbb{R}^n$.

Proof. ((a)⇒(b)) Suppose that $T(\mathbf{x}') = T(\mathbf{x}'')$. Then $A\mathbf{x}' = A\mathbf{x}''$. Multiplying on the left by A^{-1} gives $\mathbf{x}' = \mathbf{x}''$.

((b)⇒(a)) Suppose that T is one-to-one. Then the system $A\mathbf{x} = \mathbf{0}$ has unique solution $\mathbf{x} = \mathbf{0}$ in $\mathbb{R}^n$. It follows that A can be reduced by elementary row operations to the identity matrix I, and is therefore invertible.

((a)⇒(c)) For any $\mathbf{y} \in \mathbb{R}^n$, clearly $\mathbf{x} = A^{-1}\mathbf{y}$ satisfies $A\mathbf{x} = \mathbf{y}$, so that $T(\mathbf{x}) = \mathbf{y}$.

((c)⇒(a)) Suppose that $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is the standard basis for $\mathbb{R}^n$. Let $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathbb{R}^n$ be chosen to satisfy $T(\mathbf{x}_j) = \mathbf{e}_j$, so that $A\mathbf{x}_j = \mathbf{e}_j$, for every $j = 1, \ldots, n$. Write

$$C = (\, \mathbf{x}_1 \quad \ldots \quad \mathbf{x}_n \,).$$

Then $AC = I$, so that A is invertible. ◯

Definition. Suppose that the linear operator $T : \mathbb{R}^n \to \mathbb{R}^n$ has standard matrix A, where A is invertible. Then the linear operator $T^{-1} : \mathbb{R}^n \to \mathbb{R}^n$, defined by $T^{-1}(\mathbf{x}) = A^{-1}\mathbf{x}$ for every $\mathbf{x} \in \mathbb{R}^n$, is called the inverse of the linear operator T.

Remark. Clearly $T^{-1}(T(\mathbf{x})) = \mathbf{x}$ and $T(T^{-1}(\mathbf{x})) = \mathbf{x}$ for every $\mathbf{x} \in \mathbb{R}^n$.

Example 8.3.5. Consider the linear operator $T : \mathbb{R}^2 \to \mathbb{R}^2$, defined by $T(\mathbf{x}) = A\mathbf{x}$ for every $\mathbf{x} \in \mathbb{R}^2$, where

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

Clearly A is invertible, and

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

Hence the inverse linear operator is $T^{-1} : \mathbb{R}^2 \to \mathbb{R}^2$, defined by $T^{-1}(\mathbf{x}) = A^{-1}\mathbf{x}$ for every $\mathbf{x} \in \mathbb{R}^2$.

Example 8.3.6. Suppose that $T : \mathbb{R}^2 \to \mathbb{R}^2$ is anticlockwise rotation by angle θ. The reader should check that $T^{-1} : \mathbb{R}^2 \to \mathbb{R}^2$ is anticlockwise rotation by angle $2\pi - \theta$.

Next, we study the linearity properties of euclidean linear transformations which we shall use later to discuss linear transformations in arbitrary real vector spaces.

PROPOSITION 8F. *A transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is linear if and only if the following two conditions are satisfied:*
(a) For every $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, we have $T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$.
(b) For every $\mathbf{u} \in \mathbb{R}^n$ and $c \in \mathbb{R}$, we have $T(c\mathbf{u}) = cT(\mathbf{u})$.

Proof. Suppose first of all that $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation. Let A be the standard matrix for T. Then for every $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ and $c \in \mathbb{R}$, we have

$$T(\mathbf{u}+\mathbf{v}) = A(\mathbf{u}+\mathbf{v}) = A\mathbf{u} + A\mathbf{v} = T(\mathbf{u}) + T(\mathbf{v})$$

and

$$T(c\mathbf{u}) = A(c\mathbf{u}) = c(A\mathbf{u}) = cT(\mathbf{u}).$$

Suppose now that (a) and (b) hold. To show that T is linear, we need to find a matrix A such that $T(\mathbf{x}) = A\mathbf{x}$ for every $\mathbf{x} \in \mathbb{R}^n$. Suppose that $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is the standard basis for $\mathbb{R}^n$. As suggested by Proposition 8A, we write

$$A = (\, T(\mathbf{e}_1) \quad \ldots \quad T(\mathbf{e}_n) \,),$$

where $T(\mathbf{e}_j)$ is a column matrix for every $j = 1, \ldots, n$. For any vector

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

in $\mathbb{R}^n$, we have

$$A\mathbf{x} = (\, T(\mathbf{e}_1) \quad \ldots \quad T(\mathbf{e}_n) \,) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 T(\mathbf{e}_1) + \ldots + x_n T(\mathbf{e}_n).$$

Using (b) on each summand and then using (a) inductively, we obtain

$$A\mathbf{x} = T(x_1\mathbf{e}_1) + \ldots + T(x_n\mathbf{e}_n) = T(x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n) = T(\mathbf{x})$$

as required. ◯

To conclude our study of euclidean linear transformations, we briefly mention the problem of eigenvalues and eigenvectors of euclidean linear operators.

Definition. Suppose that $T : \mathbb{R}^n \to \mathbb{R}^n$ is a linear operator. Then any real number $\lambda \in \mathbb{R}$ is called an eigenvalue of T if there exists a non-zero vector $\mathbf{x} \in \mathbb{R}^n$ such that $T(\mathbf{x}) = \lambda\mathbf{x}$. This non-zero vector $\mathbf{x} \in \mathbb{R}^n$ is called an eigenvector of T corresponding to the eigenvalue λ.

Remark. Note that the equation $T(\mathbf{x}) = \lambda\mathbf{x}$ is equivalent to the equation $A\mathbf{x} = \lambda\mathbf{x}$. It follows that there is no distinction between eigenvalues and eigenvectors of T and those of the standard matrix A. We therefore do not need to discuss this problem any further.

8.4. General Linear Transformations

Suppose that V and W are real vector spaces. To define a linear transformation from V into W, we are motivated by Proposition 8F which describes the linearity properties of euclidean linear transformations.

By a transformation from V into W, we mean a function of the type $T : V \to W$, with domain V and codomain W. For every vector $\mathbf{u} \in V$, the vector $T(\mathbf{u}) \in W$ is called the image of $\mathbf{u}$ under the transformation T.

DEFINITION. A transformation $T : V \to W$ from a real vector space V into a real vector space W is called a linear transformation if the following two conditions are satisfied:
(LT1) For every $\mathbf{u}, \mathbf{v} \in V$, we have $T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$.
(LT2) For every $\mathbf{u} \in V$ and $c \in \mathbb{R}$, we have $T(c\mathbf{u}) = cT(\mathbf{u})$.

DEFINITION. A linear transformation $T : V \to V$ from a real vector space V into itself is called a linear operator on V.

EXAMPLE 8.4.1. Suppose that V and W are two real vector spaces. The transformation $T : V \to W$, where $T(\mathbf{u}) = \mathbf{0}$ for every $\mathbf{u} \in V$, is clearly linear, and is called the zero transformation from V to W.

EXAMPLE 8.4.2. Suppose that V is a real vector space. The transformation $I : V \to V$, where $I(\mathbf{u}) = \mathbf{u}$ for every $\mathbf{u} \in V$, is clearly linear, and is called the identity operator on V.

EXAMPLE 8.4.3. Suppose that V is a real vector space, and that $k \in \mathbb{R}$ is fixed. The transformation $T : V \to V$, where $T(\mathbf{u}) = k\mathbf{u}$ for every $\mathbf{u} \in V$, is clearly linear. This operator is called a dilation if $k > 1$ and a contraction if $0 < k < 1$.

EXAMPLE 8.4.4. Suppose that V is a finite dimensional vector space, with basis $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$. Define a transformation $T : V \to \mathbb{R}^n$ as follows. For every $\mathbf{u} \in V$, there exists a unique vector $(\beta_1, \ldots, \beta_n) \in \mathbb{R}^n$ such that $\mathbf{u} = \beta_1\mathbf{w}_1 + \ldots + \beta_n\mathbf{w}_n$. We let $T(\mathbf{u}) = (\beta_1, \ldots, \beta_n)$. In other words, the transformation T gives the coordinates of any vector $\mathbf{u} \in V$ with respect to the given basis $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$. Suppose now that $\mathbf{v} = \gamma_1\mathbf{w}_1 + \ldots + \gamma_n\mathbf{w}_n$ is another vector in V. Then $\mathbf{u}+\mathbf{v} = (\beta_1+\gamma_1)\mathbf{w}_1 + \ldots + (\beta_n+\gamma_n)\mathbf{w}_n$, so that

$$T(\mathbf{u}+\mathbf{v}) = (\beta_1+\gamma_1, \ldots, \beta_n+\gamma_n) = (\beta_1, \ldots, \beta_n) + (\gamma_1, \ldots, \gamma_n) = T(\mathbf{u}) + T(\mathbf{v}).$$

Also, if $c \in \mathbb{R}$, then $c\mathbf{u} = c\beta_1\mathbf{w}_1 + \ldots + c\beta_n\mathbf{w}_n$, so that

$$T(c\mathbf{u}) = (c\beta_1, \ldots, c\beta_n) = c(\beta_1, \ldots, \beta_n) = cT(\mathbf{u}).$$

Hence T is a linear transformation. We shall return to this in greater detail in the next section.

EXAMPLE 8.4.5. Suppose that P_n denotes the vector space of all polynomials with real coefficients and degree at most n. Define a transformation $T : P_n \to P_n$ as follows. For every polynomial

$$p = p_0 + p_1x + \ldots + p_nx^n$$

in P_n, we let

$$T(p) = p_n + p_{n-1}x + \ldots + p_0x^n.$$

Suppose now that $q = q_0 + q_1x + \ldots + q_nx^n$ is another polynomial in P_n. Then

$$p + q = (p_0+q_0) + (p_1+q_1)x + \ldots + (p_n+q_n)x^n,$$

so that

$$\begin{aligned} T(p+q) &= (p_n+q_n) + (p_{n-1}+q_{n-1})x + \ldots + (p_0+q_0)x^n \\ &= (p_n + p_{n-1}x + \ldots + p_0x^n) + (q_n + q_{n-1}x + \ldots + q_0x^n) = T(p) + T(q). \end{aligned}$$

Also, for any $c \in \mathbb{R}$, we have $cp = cp_0 + cp_1x + \ldots + cp_nx^n$, so that

$$T(cp) = cp_n + cp_{n-1}x + \ldots + cp_0x^n = c(p_n + p_{n-1}x + \ldots + p_0x^n) = cT(p).$$

Hence T is a linear transformation.

EXAMPLE 8.4.6. Let V denote the vector space of all real valued functions differentiable everywhere in $\mathbb{R}$, and let W denote the vector space of all real valued functions defined on $\mathbb{R}$. Consider the transformation $T : V \to W$, where $T(f) = f'$ for every $f \in V$. It is easy to check from properties of derivatives that T is a linear transformation.

EXAMPLE 8.4.7. Let V denote the vector space of all real valued functions that are Riemann integrable over the interval $[0, 1]$. Consider the transformation $T : V \to \mathbb{R}$, where

$$T(f) = \int_0^1 f(x)\,\mathrm{d}x$$

for every $f \in V$. It is easy to check from properties of the Riemann integral that T is a linear transformation.

Consider a linear transformation $T : V \to W$ from a finite dimensional real vector space V into a real vector space W. Suppose that $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a basis of V. Then every $\mathbf{u} \in V$ can be written uniquely in the form $\mathbf{u} = \beta_1\mathbf{v}_1 + \ldots + \beta_n\mathbf{v}_n$, where $\beta_1, \ldots, \beta_n \in \mathbb{R}$. It follows that

$$T(\mathbf{u}) = T(\beta_1\mathbf{v}_1 + \ldots + \beta_n\mathbf{v}_n) = T(\beta_1\mathbf{v}_1) + \ldots + T(\beta_n\mathbf{v}_n) = \beta_1T(\mathbf{v}_1) + \ldots + \beta_nT(\mathbf{v}_n).$$

We have therefore proved the following generalization of Proposition 8A.

PROPOSITION 8G. *Suppose that $T : V \to W$ is a linear transformation from a finite dimensional real vector space V into a real vector space W. Suppose further that $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a basis of V. Then T is completely determined by $T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)$.*

EXAMPLE 8.4.8. Consider a linear transformation $T : P_2 \to \mathbb{R}$, where $T(1) = 1$, $T(x) = 2$ and $T(x^2) = 3$. Since $\{1, x, x^2\}$ is a basis of P_2, this linear transformation is completely determined. In particular, we have, for example,

$$T(5 - 3x + 2x^2) = 5T(1) - 3T(x) + 2T(x^2) = 5.$$

EXAMPLE 8.4.9. Consider a linear transformation $T : \mathbb{R}^4 \to \mathbb{R}$, where $T(1, 0, 0, 0) = 1$, $T(1, 1, 0, 0) = 2$, $T(1, 1, 1, 0) = 3$ and $T(1, 1, 1, 1) = 4$. Since $\{(1, 0, 0, 0), (1, 1, 0, 0), (1, 1, 1, 0), (1, 1, 1, 1)\}$ is a basis of $\mathbb{R}^4$, this linear transformation is completely determined. In particular, we have, for example,

$$\begin{aligned} T(6, 4, 3, 1) &= T(2(1, 0, 0, 0) + (1, 1, 0, 0) + 2(1, 1, 1, 0) + (1, 1, 1, 1)) \\ &= 2T(1, 0, 0, 0) + T(1, 1, 0, 0) + 2T(1, 1, 1, 0) + T(1, 1, 1, 1) = 14. \end{aligned}$$

We also have the following generalization of Proposition 8D.

PROPOSITION 8H. *Suppose that V, W, U are real vector spaces. Suppose further that $T_1 : V \to W$ and $T_2 : W \to U$ are linear transformations. Then $T = T_2 \circ T_1 : V \to U$ is also a linear transformation.*

PROOF. Suppose that $\mathbf{u}, \mathbf{v} \in V$. Then

$$T(\mathbf{u} + \mathbf{v}) = T_2(T_1(\mathbf{u} + \mathbf{v})) = T_2(T_1(\mathbf{u}) + T_1(\mathbf{v})) = T_2(T_1(\mathbf{u})) + T_2(T_1(\mathbf{v})) = T(\mathbf{u}) + T(\mathbf{v}).$$

Also, if $c \in \mathbb{R}$, then

$$T(c\mathbf{u}) = T_2(T_1(c\mathbf{u})) = T_2(cT_1(\mathbf{u})) = cT_2(T_1(\mathbf{u})) = cT(\mathbf{u}).$$

Hence T is a linear transformation. ○

8.5. Change of Basis

Suppose that V is a real vector space, with basis $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$. Then every vector $\mathbf{u} \in V$ can be written uniquely as a linear combination

$$\mathbf{u} = \beta_1\mathbf{u}_1 + \ldots + \beta_n\mathbf{u}_n, \quad \text{where } \beta_1, \ldots, \beta_n \in \mathbb{R}. \tag{3}$$

It follows that the vector $\mathbf{u}$ can be identified with the vector $(\beta_1, \ldots, \beta_n) \in \mathbb{R}^n$.

DEFINITION. Suppose that $\mathbf{u} \in V$ and (3) holds. Then the matrix

$$[\mathbf{u}]_{\mathcal{B}} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}$$

is called the coordinate matrix of $\mathbf{u}$ relative to the basis $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$.

EXAMPLE 8.5.1. The vectors

$$\mathbf{u}_1 = (1, 2, 1, 0), \qquad \mathbf{u}_2 = (3, 3, 3, 0), \qquad \mathbf{u}_3 = (2, -10, 0, 0), \qquad \mathbf{u}_4 = (-2, 1, -6, 2)$$

are linearly independent in $\mathbb{R}^4$, and so $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$ is a basis of $\mathbb{R}^4$. It follows that for any $\mathbf{u} = (x, y, z, w) \in \mathbb{R}^4$, we can write

$$\mathbf{u} = \beta_1\mathbf{u}_1 + \beta_2\mathbf{u}_2 + \beta_3\mathbf{u}_3 + \beta_4\mathbf{u}_4.$$

In matrix notation, this becomes

$$\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 1 & 3 & 2 & -2 \\ 2 & 3 & -10 & 1 \\ 1 & 3 & 0 & -6 \\ 0 & 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{pmatrix},$$

so that

$$[\mathbf{u}]_{\mathcal{B}} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{pmatrix} = \begin{pmatrix} 1 & 3 & 2 & -2 \\ 2 & 3 & -10 & 1 \\ 1 & 3 & 0 & -6 \\ 0 & 0 & 0 & 2 \end{pmatrix}^{-1} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix}.$$

REMARK. Consider a function $\phi : V \to \mathbb{R}^n$, where $\phi(\mathbf{u}) = [\mathbf{u}]_{\mathcal{B}}$ for every $\mathbf{u} \in V$. It is not difficult to see that this function gives rise to a one-to-one correspondence between the elements of V and the elements of $\mathbb{R}^n$. Furthermore, note that

$$[\mathbf{u} + \mathbf{v}]_{\mathcal{B}} = [\mathbf{u}]_{\mathcal{B}} + [\mathbf{v}]_{\mathcal{B}} \qquad \text{and} \qquad [c\mathbf{u}]_{\mathcal{B}} = c[\mathbf{u}]_{\mathcal{B}},$$

so that $\phi(\mathbf{u} + \mathbf{v}) = \phi(\mathbf{u}) + \phi(\mathbf{v})$ and $\phi(c\mathbf{u}) = c\phi(\mathbf{u})$ for every $\mathbf{u}, \mathbf{v} \in V$ and $c \in \mathbb{R}$. Thus ϕ is a linear transformation, and preserves much of the structure of V. We also say that V is isomorphic to $\mathbb{R}^n$. In practice, once we have made this identification between vectors and their coordinate matrices, then we can basically forget about the basis $\mathcal{B}$ and imagine that we are working in $\mathbb{R}^n$ with the standard basis.

Clearly, if we change from one basis $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ to another basis $\mathcal{C} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ of V, then we also need to find a way of calculating $[\mathbf{u}]_{\mathcal{C}}$ in terms of $[\mathbf{u}]_{\mathcal{B}}$ for every vector $\mathbf{u} \in V$. To do this, note that each of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ can be written uniquely as a linear combination of the vectors $\mathbf{u}_1, \ldots, \mathbf{u}_n$. Suppose that for $i = 1, \ldots, n$, we have

$$\mathbf{v}_i = a_{1i}\mathbf{u}_1 + \ldots + a_{ni}\mathbf{u}_n, \quad \text{where } a_{1i}, \ldots, a_{ni} \in \mathbb{R},$$

so that

$$[\mathbf{v}_i]_{\mathcal{B}} = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{ni} \end{pmatrix}.$$

For every $\mathbf{u} \in V$, we can write

$$\mathbf{u} = \beta_1 \mathbf{u}_1 + \ldots + \beta_n \mathbf{u}_n = \gamma_1 \mathbf{v}_1 + \ldots + \gamma_n \mathbf{v}_n, \quad \text{where } \beta_1, \ldots, \beta_n, \gamma_1, \ldots, \gamma_n \in \mathbb{R},$$

so that

$$[\mathbf{u}]_{\mathcal{B}} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \qquad \text{and} \qquad [\mathbf{u}]_{\mathcal{C}} = \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{pmatrix}.$$

Clearly

$$\begin{aligned} \mathbf{u} &= \gamma_1 \mathbf{v}_1 + \ldots + \gamma_n \mathbf{v}_n \\ &= \gamma_1 (a_{11} \mathbf{u}_1 + \ldots + a_{n1} \mathbf{u}_n) + \ldots + \gamma_n (a_{1n} \mathbf{u}_1 + \ldots + a_{nn} \mathbf{u}_n) \\ &= (\gamma_1 a_{11} + \ldots + \gamma_n a_{1n}) \mathbf{u}_1 + \ldots + (\gamma_1 a_{n1} + \ldots + \gamma_n a_{nn}) \mathbf{u}_n \\ &= \beta_1 \mathbf{u}_1 + \ldots + \beta_n \mathbf{u}_n. \end{aligned}$$

Hence

$$\begin{aligned} \beta_1 &= \gamma_1 a_{11} + \ldots + \gamma_n a_{1n}, \\ &\vdots \\ \beta_n &= \gamma_1 a_{n1} + \ldots + \gamma_n a_{nn}. \end{aligned}$$

Written in matrix notation, we have

$$\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{pmatrix}.$$

We have proved the following result.

PROPOSITION 8J. *Suppose that $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ and $\mathcal{C} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ are two bases of a real vector space V. Then for every $\mathbf{u} \in V$, we have*

$$[\mathbf{u}]_{\mathcal{B}} = P[\mathbf{u}]_{\mathcal{C}},$$

where the columns of the matrix

$$P = (\, [\mathbf{v}_1]_{\mathcal{B}} \quad \ldots \quad [\mathbf{v}_n]_{\mathcal{B}} \,)$$

are precisely the coordinate matrices of the elements of $\mathcal{C}$ relative to the basis $\mathcal{B}$.

Remark. Strictly speaking, Proposition 8J gives $[\mathbf{u}]_{\mathcal{B}}$ in terms of $[\mathbf{u}]_{\mathcal{C}}$. However, note that the matrix P is invertible (why?), so that $[\mathbf{u}]_{\mathcal{C}} = P^{-1}[\mathbf{u}]_{\mathcal{B}}$.

Definition. The matrix P in Proposition 8J is sometimes called the transition matrix from the basis $\mathcal{C}$ to the basis $\mathcal{B}$.

Example 8.5.2. We know that with

$$\mathbf{u}_1 = (1,2,1,0), \qquad \mathbf{u}_2 = (3,3,3,0), \qquad \mathbf{u}_3 = (2,-10,0,0), \qquad \mathbf{u}_4 = (-2,1,-6,2),$$

and with

$$\mathbf{v}_1 = (1,2,1,0), \qquad \mathbf{v}_2 = (1,-1,1,0), \qquad \mathbf{v}_3 = (1,0,-1,0), \qquad \mathbf{v}_4 = (0,0,0,2),$$

both $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$ and $\mathcal{C} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ are bases of $\mathbb{R}^4$. It is easy to check that

$$\begin{aligned} \mathbf{v}_1 &= \mathbf{u}_1, \\ \mathbf{v}_2 &= -2\mathbf{u}_1 + \mathbf{u}_2, \\ \mathbf{v}_3 &= 11\mathbf{u}_1 - 4\mathbf{u}_2 + \mathbf{u}_3, \\ \mathbf{v}_4 &= -27\mathbf{u}_1 + 11\mathbf{u}_2 - 2\mathbf{u}_3 + \mathbf{u}_4, \end{aligned}$$

so that

$$P = (\, [\mathbf{v}_1]_{\mathcal{B}} \quad [\mathbf{v}_2]_{\mathcal{B}} \quad [\mathbf{v}_3]_{\mathcal{B}} \quad [\mathbf{v}_4]_{\mathcal{B}} \,) = \begin{pmatrix} 1 & -2 & 11 & -27 \\ 0 & 1 & -4 & 11 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Hence $[\mathbf{u}]_{\mathcal{B}} = P[\mathbf{u}]_{\mathcal{C}}$ for every $\mathbf{u} \in \mathbb{R}^4$. It is also easy to check that

$$\begin{aligned} \mathbf{u}_1 &= \mathbf{v}_1, \\ \mathbf{u}_2 &= 2\mathbf{v}_1 + \mathbf{v}_2, \\ \mathbf{u}_3 &= -3\mathbf{v}_1 + 4\mathbf{v}_2 + \mathbf{v}_3, \\ \mathbf{u}_4 &= -\mathbf{v}_1 - 3\mathbf{v}_2 + 2\mathbf{v}_3 + \mathbf{v}_4, \end{aligned}$$

so that

$$Q = (\, [\mathbf{u}_1]_{\mathcal{C}} \quad [\mathbf{u}_2]_{\mathcal{C}} \quad [\mathbf{u}_3]_{\mathcal{C}} \quad [\mathbf{u}_4]_{\mathcal{C}} \,) = \begin{pmatrix} 1 & 2 & -3 & -1 \\ 0 & 1 & 4 & -3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Hence $[\mathbf{u}]_{\mathcal{C}} = Q[\mathbf{u}]_{\mathcal{B}}$ for every $\mathbf{u} \in \mathbb{R}^4$. Note that $PQ = I$. Now let $\mathbf{u} = (6,-1,2,2)$. We can check that $\mathbf{u} = \mathbf{v}_1 + 3\mathbf{v}_2 + 2\mathbf{v}_3 + \mathbf{v}_4$, so that

$$[\mathbf{u}]_{\mathcal{C}} = \begin{pmatrix} 1 \\ 3 \\ 2 \\ 1 \end{pmatrix}.$$

Then

$$[\mathbf{u}]_{\mathcal{B}} = \begin{pmatrix} 1 & -2 & 11 & -27 \\ 0 & 1 & -4 & 11 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} -10 \\ 6 \\ 0 \\ 1 \end{pmatrix}.$$

Check that $\mathbf{u} = -10\mathbf{u}_1 + 6\mathbf{u}_2 + \mathbf{u}_4$.

EXAMPLE 8.5.3. Consider the vector space P_2. It is not too difficult to check that

$$u_1 = 1 + x, \qquad u_2 = 1 + x^2, \qquad u_3 = x + x^2$$

form a basis of P_2. Let $u = 1 + 4x - x^2$. Then $u = \beta_1 u_1 + \beta_2 u_2 + \beta_3 u_3$, where

$$1 + 4x - x^2 = \beta_1(1 + x) + \beta_2(1 + x^2) + \beta_3(x + x^2) = (\beta_1 + \beta_2) + (\beta_1 + \beta_3)x + (\beta_2 + \beta_3)x^2,$$

so that $\beta_1 + \beta_2 = 1$, $\beta_1 + \beta_3 = 4$ and $\beta_2 + \beta_3 = -1$. Hence $(\beta_1, \beta_2, \beta_3) = (3, -2, 1)$. If we write $\mathcal{B} = \{u_1, u_2, u_3\}$, then

$$[u]_{\mathcal{B}} = \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix}.$$

On the other hand, it is also not too difficult to check that

$$v_1 = 1, \qquad v_2 = 1 + x, \qquad v_3 = 1 + x + x^2$$

form a basis of P_2. Also $u = \gamma_1 v_1 + \gamma_2 v_2 + \gamma_3 v_3$, where

$$1 + 4x - x^2 = \gamma_1 + \gamma_2(1 + x) + \gamma_3(1 + x + x^2) = (\gamma_1 + \gamma_2 + \gamma_3) + (\gamma_2 + \gamma_3)x + \gamma_3 x^2,$$

so that $\gamma_1 + \gamma_2 + \gamma_3 = 1$, $\gamma_2 + \gamma_3 = 4$ and $\gamma_3 = -1$. Hence $(\gamma_1, \gamma_2, \gamma_3) = (-3, 5, -1)$. If we write $\mathcal{C} = \{v_1, v_2, v_3\}$, then

$$[u]_{\mathcal{C}} = \begin{pmatrix} -3 \\ 5 \\ -1 \end{pmatrix}.$$

Next, note that

$$\begin{aligned} v_1 &= \tfrac{1}{2}u_1 + \tfrac{1}{2}u_2 - \tfrac{1}{2}u_3, \\ v_2 &= u_1, \\ v_3 &= \tfrac{1}{2}u_1 + \tfrac{1}{2}u_2 + \tfrac{1}{2}u_3. \end{aligned}$$

Hence

$$P = (\, [v_1]_{\mathcal{B}} \quad [v_2]_{\mathcal{B}} \quad [v_3]_{\mathcal{B}} \,) = \begin{pmatrix} 1/2 & 1 & 1/2 \\ 1/2 & 0 & 1/2 \\ -1/2 & 0 & 1/2 \end{pmatrix}.$$

To verify that $[u]_{\mathcal{B}} = P[u]_{\mathcal{C}}$, note that

$$\begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1/2 & 1 & 1/2 \\ 1/2 & 0 & 1/2 \\ -1/2 & 0 & 1/2 \end{pmatrix} \begin{pmatrix} -3 \\ 5 \\ -1 \end{pmatrix}.$$

8.6. Kernel and Range

Consider first of all a euclidean linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$. Suppose that A is the standard matrix for T. Then the range of the transformation T is given by

$$R(T) = \{T(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^n\} = \{A\mathbf{x} : \mathbf{x} \in \mathbb{R}^n\}.$$

It follows that $R(T)$ is the set of all linear combinations of the columns of the matrix A, and is therefore the column space of A. On the other hand, the set

$$\{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{0}\}$$

is the nullspace of A.

Recall that the sum of the dimension of the nullspace of A and dimension of the column space of A is equal to the number of columns of A. This is known as the Rank-nullity theorem. The purpose of this section is to extend this result to the setting of linear transformations. To do this, we need the following generalization of the idea of the nullspace and the column space.

DEFINITION. Suppose that $T : V \to W$ is a linear transformation from a real vector space V into a real vector space W. Then the set

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

is called the kernel of T, and the set

$$R(T) = \{T(\mathbf{u}) : \mathbf{u} \in V\}$$

is called the range of T.

EXAMPLE 8.6.1. For a euclidean linear transformation T with standard matrix A, we have shown that $\ker(T)$ is the nullspace of A, while $R(T)$ is the column space of A.

EXAMPLE 8.6.2. Suppose that $T : V \to W$ is the zero transformation. Clearly we have $\ker(T) = V$ and $R(T) = \{\mathbf{0}\}$.

EXAMPLE 8.6.3. Suppose that $T : V \to V$ is the identity operator on V. Clearly we have $\ker(T) = \{\mathbf{0}\}$ and $R(T) = V$.

EXAMPLE 8.6.4. Suppose that $T : \mathbb{R}^2 \to \mathbb{R}^2$ is orthogonal projection onto the x_1-axis. Then $\ker(T)$ is the x_2-axis, while $R(T)$ is the x_1-axis.

EXAMPLE 8.6.5. Suppose that $T : \mathbb{R}^n \to \mathbb{R}^n$ is one-to-one. Then $\ker(T) = \{\mathbf{0}\}$ and $R(T) = \mathbb{R}^n$, in view of Proposition 8E.

EXAMPLE 8.6.6. Consider the linear transformation $T : V \to W$, where V denotes the vector space of all real valued functions differentiable everywhere in $\mathbb{R}$, where W denotes the space of all real valued functions defined in $\mathbb{R}$, and where $T(f) = f'$ for every $f \in V$. Then $\ker(T)$ is the set of all differentiable functions with derivative 0, and so is the set of all constant functions in $\mathbb{R}$.

EXAMPLE 8.6.7. Consider the linear transformation $T : V \to \mathbb{R}$, where V denotes the vector space of all real valued functions Riemann integrable over the interval $[0, 1]$, and where

$$T(f) = \int_0^1 f(x)\,\mathrm{d}x$$

for every $f \in V$. Then $\ker(T)$ is the set of all Riemann integrable functions in $[0, 1]$ with zero mean, while $R(T) = \mathbb{R}$.

PROPOSITION 8K. *Suppose that $T : V \to W$ is a linear transformation from a real vector space V into a real vector space W. Then $\ker(T)$ is a subspace of V, while $R(T)$ is a subspace of W.*

PROOF. Since $T(\mathbf{0}) = \mathbf{0}$, it follows that $\mathbf{0} \in \ker(T) \subseteq V$ and $\mathbf{0} \in R(T) \subseteq W$. For any $\mathbf{u}, \mathbf{v} \in \ker(T)$, we have

$$T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v}) = \mathbf{0} + \mathbf{0} = \mathbf{0},$$

so that $\mathbf{u} + \mathbf{v} \in \ker(T)$. Suppose further that $c \in \mathbb{R}$. Then

$$T(c\mathbf{u}) = cT(\mathbf{u}) = c\mathbf{0} = \mathbf{0},$$

so that $c\mathbf{u} \in \ker(T)$. Hence $\ker(T)$ is a subspace of V. Suppose next that $\mathbf{w}, \mathbf{z} \in R(T)$. Then there exist $\mathbf{u}, \mathbf{v} \in V$ such that $T(\mathbf{u}) = \mathbf{w}$ and $T(\mathbf{v}) = \mathbf{z}$. Hence

$$T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v}) = \mathbf{w} + \mathbf{z},$$

so that $\mathbf{w} + \mathbf{z} \in R(T)$. Suppose further that $c \in \mathbb{R}$. Then

$$T(c\mathbf{u}) = cT(\mathbf{u}) = c\mathbf{w},$$

so that $c\mathbf{w} \in R(T)$. Hence $R(T)$ is a subspace of W. ○

To complete this section, we prove the following generalization of the Rank-nullity theorem.

PROPOSITION 8L. *Suppose that $T : V \to W$ is a linear transformation from an n-dimensional real vector space V into a real vector space W. Then*

$$\dim \ker(T) + \dim R(T) = n.$$

PROOF. Suppose first of all that $\dim \ker(T) = n$. Then $\ker(T) = V$, and so $R(T) = \{\mathbf{0}\}$, and the result follows immediately. Suppose next that $\dim \ker(T) = 0$, so that $\ker(T) = \{\mathbf{0}\}$. If $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a basis of V, then it follows that $T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)$ are linearly independent in W, for otherwise there exist $c_1, \ldots, c_n \in \mathbb{R}$, not all zero, such that

$$c_1T(\mathbf{v}_1) + \ldots + c_nT(\mathbf{v}_n) = \mathbf{0},$$

so that $T(c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n) = \mathbf{0}$, a contradiction since $c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n \neq \mathbf{0}$. On the other hand, elements of $R(T)$ are linear combinations of $T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)$. Hence $\dim R(T) = n$, and the result again follows immediately. We may therefore assume that $\dim \ker(T) = r$, where $1 \leq r < n$. Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ be a basis of $\ker(T)$. This basis can be extended to a basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_r, \mathbf{v}_{r+1}, \ldots, \mathbf{v}_n\}$ of V. It suffices to show that

$$\{T(\mathbf{v}_{r+1}), \ldots, T(\mathbf{v}_n)\} \tag{4}$$

is a basis of $R(T)$. Suppose that $\mathbf{u} \in V$. Then there exist $\beta_1, \ldots, \beta_n \in \mathbb{R}$ such that

$$\mathbf{u} = \beta_1\mathbf{v}_1 + \ldots + \beta_r\mathbf{v}_r + \beta_{r+1}\mathbf{v}_{r+1} + \ldots + \beta_n\mathbf{v}_n,$$

so that

$$\begin{aligned} T(\mathbf{u}) &= \beta_1T(\mathbf{v}_1) + \ldots + \beta_rT(\mathbf{v}_r) + \beta_{r+1}T(\mathbf{v}_{r+1}) + \ldots + \beta_nT(\mathbf{v}_n) \\ &= \beta_{r+1}T(\mathbf{v}_{r+1}) + \ldots + \beta_nT(\mathbf{v}_n). \end{aligned}$$

It follows that (4) spans $R(T)$. It remains to prove that its elements are linearly independent. Suppose that $c_{r+1}, \ldots, c_n \in \mathbb{R}$ and

$$c_{r+1}T(\mathbf{v}_{r+1}) + \ldots + c_nT(\mathbf{v}_n) = \mathbf{0}. \tag{5}$$

We need to show that

$$c_{r+1} = \ldots = c_n = 0. \tag{6}$$

By linearity, it follows from (5) that $T(c_{r+1}\mathbf{v}_{r+1} + \ldots + c_n\mathbf{v}_n) = \mathbf{0}$, so that

$$c_{r+1}\mathbf{v}_{r+1} + \ldots + c_n\mathbf{v}_n \in \ker(T).$$

Hence there exist $c_1, \ldots, c_r \in \mathbb{R}$ such that

$$c_{r+1}\mathbf{v}_{r+1} + \ldots + c_n\mathbf{v}_n = c_1\mathbf{v}_1 + \ldots + c_r\mathbf{v}_r,$$

so that

$$c_1\mathbf{v}_1 + \ldots + c_r\mathbf{v}_r - c_{r+1}\mathbf{v}_{r+1} - \ldots - c_n\mathbf{v}_n = \mathbf{0}.$$

Since $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a basis of V, it follows that $c_1 = \ldots = c_r = c_{r+1} = \ldots = c_n = 0$, so that (6) holds. This completes the proof. ◯

REMARK. We sometimes say that $\dim R(T)$ and $\dim \ker(T)$ are respectively the rank and the nullity of the linear transformation T.

8.7. Inverse Linear Transformations

In this section, we generalize some of the ideas first discussed in Section 8.3.

DEFINITION. A linear transformation $T : V \to W$ from a real vector space V into a real vector space W is said to be one-to-one if for every $\mathbf{u}', \mathbf{u}'' \in V$, we have $\mathbf{u}' = \mathbf{u}''$ whenever $T(\mathbf{u}') = T(\mathbf{u}'')$.

The result below follows immediately from our definition.

PROPOSITION 8M. *Suppose that $T : V \to W$ is a linear transformation from a real vector space V into a real vector space W. Then T is one-to-one if and only if* $\ker(T) = \{\mathbf{0}\}$.

PROOF. (⇒) Clearly $\mathbf{0} \in \ker(T)$. Suppose that $\ker(T) \neq \{\mathbf{0}\}$. Then there exists a non-zero $\mathbf{v} \in \ker(T)$. It follows that $T(\mathbf{v}) = T(\mathbf{0})$, and so T is not one-to-one.

(⇐) Suppose that $\ker(T) = \{\mathbf{0}\}$. Given any $\mathbf{u}', \mathbf{u}'' \in V$, we have

$$T(\mathbf{u}') - T(\mathbf{u}'') = T(\mathbf{u}' - \mathbf{u}'') = \mathbf{0}$$

if and only if $\mathbf{u}' - \mathbf{u}'' = \mathbf{0}$; in other words, if and only if $\mathbf{u}' = \mathbf{u}''$. ◯

We have the following generalization of Proposition 8E.

PROPOSITION 8N. *Suppose that $T : V \to V$ is a linear operator on a finite-dimensional real vector space V. Then the following statements are equivalent:*
(a) The linear operator T is one-to-one.
(b) We have $\ker(T) = \{\mathbf{0}\}$.
(c) The range of T is V; in other words, $R(T) = V$.

PROOF. The equivalence of (a) and (b) is established by Proposition 8M. The equivalence of (b) and (c) follows from Proposition 8L. ◯

Suppose that $T : V \to W$ is a one-to-one linear transformation from a real vector space V into a real vector space W. Then for every $\mathbf{w} \in R(T)$, there exists exactly one $\mathbf{u} \in V$ such that $T(\mathbf{u}) = \mathbf{w}$. We can therefore define a transformation $T^{-1} : R(T) \to V$ by writing $T^{-1}(\mathbf{w}) = \mathbf{u}$, where $\mathbf{u} \in V$ is the unique vector satisfying $T(\mathbf{u}) = \mathbf{w}$.

PROPOSITION 8P. *Suppose that $T : V \to W$ is a one-to-one linear transformation from a real vector space V into a real vector space W. Then $T^{-1} : R(T) \to V$ is a linear transformation.*

PROOF. Suppose that $\mathbf{w}, \mathbf{z} \in R(T)$. Then there exist $\mathbf{u}, \mathbf{v} \in V$ such that $T^{-1}(\mathbf{w}) = \mathbf{u}$ and $T^{-1}(\mathbf{z}) = \mathbf{v}$. It follows that $T(\mathbf{u}) = \mathbf{w}$ and $T(\mathbf{v}) = \mathbf{z}$, so that $T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v}) = \mathbf{w}+\mathbf{z}$, whence

$$T^{-1}(\mathbf{w}+\mathbf{z}) = \mathbf{u}+\mathbf{v} = T^{-1}(\mathbf{w}) + T^{-1}(\mathbf{z}).$$

Suppose further that $c \in \mathbb{R}$. Then $T(c\mathbf{u}) = c\mathbf{w}$, so that

$$T^{-1}(c\mathbf{w}) = c\mathbf{u} = cT^{-1}(\mathbf{w}).$$

This completes the proof. ○

We also have the following result concerning compositions of linear transformations and which requires no further proof, in view of our knowledge concerning inverse functions.

PROPOSITION 8Q. *Suppose that V, W, U are real vector spaces. Suppose further that $T_1 : V \to W$ and $T_2 : W \to U$ are one-to-one linear transformations. Then*
(a) the linear transformation $T_2 \circ T_1 : V \to U$ is one-to-one; and
(b) $(T_2 \circ T_1)^{-1} = T_1^{-1} \circ T_2^{-1}$.

8.8. Matrices of General Linear Transformations

Suppose that $T : V \to W$ is a linear transformation from a real vector space V to a real vector space W. Suppose further that the vector spaces V and W are finite dimensional, with $\dim V = n$ and $\dim W = m$. We shall show that if we make use of a basis $\mathcal{B}$ of V and a basis $\mathcal{C}$ of W, then it is possible to describe T indirectly in terms of some matrix A. The main idea is to make use of coordinate matrices relative to the bases $\mathcal{B}$ and $\mathcal{C}$.

Let us recall some discussion in Section 8.5. Suppose that $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a basis of V. Then every vector $\mathbf{v} \in V$ can be written uniquely as a linear combination

$$\mathbf{v} = \beta_1\mathbf{v}_1 + \ldots + \beta_n\mathbf{v}_n, \quad \text{where } \beta_1, \ldots, \beta_n \in \mathbb{R}. \tag{7}$$

The matrix

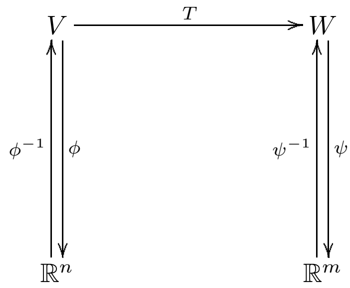
$$[\mathbf{v}]_{\mathcal{B}} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \tag{8}$$

is the coordinate matrix of $\mathbf{v}$ relative to the basis $\mathcal{B}$.

Consider now a transformation $\phi : V \to \mathbb{R}^n$, where $\phi(\mathbf{v}) = [\mathbf{v}]_{\mathcal{B}}$ for every $\mathbf{v} \in V$. The proof of the following result is straightforward.

PROPOSITION 8R. *Suppose that the real vector space V has basis $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. Then the transformation $\phi : V \to \mathbb{R}^n$, where $\phi(\mathbf{v}) = [\mathbf{v}]_{\mathcal{B}}$ satisfies (7) and (8) for every $\mathbf{v} \in V$, is a one-to-one linear transformation, with range $R(\phi) = \mathbb{R}^n$. Furthermore, the inverse linear transformation $\phi^{-1} : \mathbb{R}^n \to V$ is also one-to-one, with range $R(\phi^{-1}) = V$.*

Suppose next that $\mathcal{C} = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ is a basis of W. Then we can define a linear transformation $\psi : W \to \mathbb{R}^m$, where $\psi(\mathbf{w}) = [\mathbf{w}]_{\mathcal{C}}$ for every $\mathbf{w} \in W$, in a similar way. We now have the following diagram of linear transformations.



Clearly the composition

$$S = \psi \circ T \circ \phi^{-1} : \mathbb{R}^n \to \mathbb{R}^m$$

is a euclidean linear transformation, and can therefore be described in terms of a standard matrix A. Our task is to determine this matrix A in terms of T and the bases $\mathcal{B}$ and $\mathcal{C}$.

We know from Proposition 8A that

$$A = (\, S(\mathbf{e}_1) \quad \ldots \quad S(\mathbf{e}_n) \,),$$

where $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is the standard basis for $\mathbb{R}^n$. For every $j = 1, \ldots, n$, we have

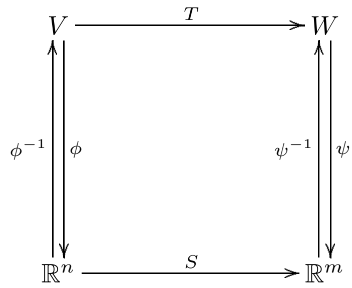
$$S(\mathbf{e}_j) = (\psi \circ T \circ \phi^{-1})(\mathbf{e}_j) = \psi(T(\phi^{-1}(\mathbf{e}_j))) = \psi(T(\mathbf{v}_j)) = [T(\mathbf{v}_j)]_{\mathcal{C}}.$$

It follows that

$$A = (\, [T(\mathbf{v}_1)]_{\mathcal{C}} \quad \ldots \quad [T(\mathbf{v}_n)]_{\mathcal{C}} \,). \tag{9}$$

Definition. The matrix A given by (9) is called the matrix for the linear transformation T with respect to the bases $\mathcal{B}$ and $\mathcal{C}$.

We now have the following diagram of linear transformations.



Hence we can write T as the composition

$$T = \psi^{-1} \circ S \circ \phi : V \to W.$$

For every $\mathbf{v} \in V$, we have the following:

$$\mathbf{v} \xrightarrow{\ \phi\ } [\mathbf{v}]_{\mathcal{B}} \xrightarrow{\ S\ } A[\mathbf{v}]_{\mathcal{B}} \xrightarrow{\ \psi^{-1}\ } \psi^{-1}(A[\mathbf{v}]_{\mathcal{B}})$$

More precisely, if $\mathbf{v} = \beta_1\mathbf{v}_1 + \ldots + \beta_n\mathbf{v}_n$, then

$$[\mathbf{v}]_{\mathcal{B}} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \qquad \text{and} \qquad A[\mathbf{v}]_{\mathcal{B}} = A\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_m \end{pmatrix},$$

say, and so $T(\mathbf{v}) = \psi^{-1}(A[\mathbf{v}]_{\mathcal{B}}) = \gamma_1\mathbf{w}_1 + \ldots + \gamma_m\mathbf{w}_m$. We have proved the following result.

PROPOSITION 8S. *Suppose that $T : V \to W$ is a linear transformation from a real vector space V into a real vector space W. Suppose further that V and W are finite dimensional, with bases $\mathcal{B}$ and $\mathcal{C}$ respectively, and that A is the matrix for the linear transformation T with respect to the bases $\mathcal{B}$ and $\mathcal{C}$. Then for every $\mathbf{v} \in V$, we have $T(\mathbf{v}) = \mathbf{w}$, where $\mathbf{w} \in W$ is the unique vector satisfying $[\mathbf{w}]_{\mathcal{C}} = A[\mathbf{v}]_{\mathcal{B}}$.*

REMARK. In the special case when $V = W$, the linear transformation $T : V \to W$ is a linear operator on T. Of course, we may choose a basis $\mathcal{B}$ for the domain V of T and a basis $\mathcal{C}$ for the codomain V of T. In the case when T is the identity linear operator, we often choose $\mathcal{B} \neq \mathcal{C}$ since this represents a change of basis. In the case when T is not the identity operator, we often choose $\mathcal{B} = \mathcal{C}$ for the sake of convenience; we then say that A is the matrix for the linear operator T with respect to the basis $\mathcal{B}$.

EXAMPLE 8.8.1. Consider an operator $T : P_3 \to P_3$ on the real vector space P_3 of all polynomials with real coefficients and degree at most 3, where for every polynomial $p(x)$ in P_3, we have $T(p(x)) = xp'(x)$, the product of x with the formal derivative $p'(x)$ of $p(x)$. The reader is invited to check that T is a linear operator. Now consider the basis $\mathcal{B} = \{1, x, x^2, x^3\}$ of P_3. The matrix for T with respect to $\mathcal{B}$ is given by

$$A = (\,[T(1)]_{\mathcal{B}} \quad [T(x)]_{\mathcal{B}} \quad [T(x^2)]_{\mathcal{B}} \quad [T(x^3)]_{\mathcal{B}}\,) = (\,[0]_{\mathcal{B}} \quad [x]_{\mathcal{B}} \quad [2x^2]_{\mathcal{B}} \quad [3x^3]_{\mathcal{B}}\,) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Suppose that $p(x) = 1 + 2x + 4x^2 + 3x^3$. Then

$$[p(x)]_{\mathcal{B}} = \begin{pmatrix} 1 \\ 2 \\ 4 \\ 3 \end{pmatrix} \qquad \text{and} \qquad A[p(x)]_{\mathcal{B}} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \\ 4 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 8 \\ 9 \end{pmatrix},$$

so that $T(p(x)) = 2x + 8x^2 + 9x^3$. This can be easily verified by noting that

$$T(p(x)) = xp'(x) = x(2 + 8x + 9x^2) = 2x + 8x^2 + 9x^3.$$

In general, if $p(x) = p_0 + p_1x + p_2x^2 + p_3x^3$, then

$$[p(x)]_{\mathcal{B}} = \begin{pmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \end{pmatrix} \qquad \text{and} \qquad A[p(x)]_{\mathcal{B}} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}\begin{pmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \end{pmatrix} = \begin{pmatrix} 0 \\ p_1 \\ 2p_2 \\ 3p_3 \end{pmatrix},$$

so that $T(p(x)) = p_1x + 2p_2x^2 + 3p_3x^3$. Observe that

$$T(p(x)) = xp'(x) = x(p_1 + 2p_2x + 3p_3x^2) = p_1x + 2p_2x^2 + 3p_3x^3,$$

verifying our result.

EXAMPLE 8.8.2. Consider the linear operator $T : \mathbb{R}^2 \to \mathbb{R}^2$, given by $T(x_1, x_2) = (2x_1 + x_2, x_1 + 3x_2)$ for every $(x_1, x_2) \in \mathbb{R}^2$. Consider also the basis $\mathcal{B} = \{(1, 0), (1, 1)\}$ of $\mathbb{R}^2$. Then the matrix for T with respect to $\mathcal{B}$ is given by

$$A = (\, [T(1,0)]_{\mathcal{B}} \quad [T(1,1)]_{\mathcal{B}} \,) = (\, [(2,1)]_{\mathcal{B}} \quad [(3,4)]_{\mathcal{B}} \,) = \begin{pmatrix} 1 & -1 \\ 1 & 4 \end{pmatrix}.$$

Suppose that $(x_1, x_2) = (3, 2)$. Then

$$[(3,2)]_{\mathcal{B}} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \qquad \text{and} \qquad A[(3,2)]_{\mathcal{B}} = \begin{pmatrix} 1 & -1 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} -1 \\ 9 \end{pmatrix},$$

so that $T(3, 2) = -(1, 0) + 9(1, 1) = (8, 9)$. This can be easily verified directly. In general, we have

$$[(x_1,x_2)]_{\mathcal{B}} = \begin{pmatrix} x_1 - x_2 \\ x_2 \end{pmatrix} \qquad \text{and} \qquad A[(x_1,x_2)]_{\mathcal{B}} = \begin{pmatrix} 1 & -1 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} x_1 - x_2 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 - 2x_2 \\ x_1 + 3x_2 \end{pmatrix},$$

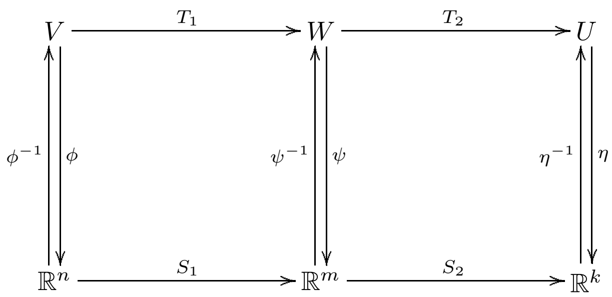
so that $T(x_1, x_2) = (x_1 - 2x_2)(1, 0) + (x_1 + 3x_2)(1, 1) = (2x_1 + x_2, x_1 + 3x_2)$.

EXAMPLE 8.8.3. Suppose that $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation. Suppose further that $\mathcal{B}$ and $\mathcal{C}$ are the standard bases for $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively. Then the matrix for T with respect to $\mathcal{B}$ and $\mathcal{C}$ is given by

$$A = (\, [T(\mathbf{e}_1)]_{\mathcal{C}} \quad \ldots \quad [T(\mathbf{e}_n)]_{\mathcal{C}} \,) = (\, T(\mathbf{e}_1) \quad \ldots \quad T(\mathbf{e}_n) \,),$$

so it follows from Proposition 8A that A is simply the standard matrix for T.

Suppose now that $T_1 : V \to W$ and $T_2 : W \to U$ are linear transformations, where the real vector spaces V, W, U are finite dimensional, with respective bases $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, $\mathcal{C} = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ and $\mathcal{D} = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$. We then have the following diagram of linear transformations.



Here $\eta : U \to \mathbb{R}^k$, where $\eta(\mathbf{u}) = [\mathbf{u}]_{\mathcal{D}}$ for every $\mathbf{u} \in U$, is a linear transformation, and

$$S_1 = \psi \circ T_1 \circ \phi^{-1} : \mathbb{R}^n \to \mathbb{R}^m \qquad \text{and} \qquad S_2 = \eta \circ T_2 \circ \psi^{-1} : \mathbb{R}^m \to \mathbb{R}^k$$

are euclidean linear transformations. Suppose that A_1 and A_2 are respectively the standard matrices for S_1 and S_2, so that they are respectively the matrix for T_1 with respect to $\mathcal{B}$ and $\mathcal{C}$ and the matrix for T_2 with respect to $\mathcal{C}$ and $\mathcal{D}$. Clearly

$$S_2 \circ S_1 = \eta \circ T_2 \circ T_1 \circ \phi^{-1} : \mathbb{R}^n \to \mathbb{R}^k.$$

It follows that $A_2 A_1$ is the standard matrix for $S_2 \circ S_1$, and so is the matrix for $T_2 \circ T_1$ with respect to the bases $\mathcal{B}$ and $\mathcal{D}$. To summarize, we have the following result.

PROPOSITION 8T. *Suppose that $T_1 : V \to W$ and $T_2 : W \to U$ are linear transformations, where the real vector spaces V, W, U are finite dimensional, with bases $\mathcal{B}, \mathcal{C}, \mathcal{D}$ respectively. Suppose further that A_1 is the matrix for the linear transformation T_1 with respect to the bases $\mathcal{B}$ and $\mathcal{C}$, and that A_2 is the matrix for the linear transformation T_2 with respect to the bases $\mathcal{C}$ and $\mathcal{D}$. Then A_2A_1 is the matrix for the linear transformation $T_2 \circ T_1$ with respect to the bases $\mathcal{B}$ and $\mathcal{D}$.*

EXAMPLE 8.8.4. Consider the linear operator $T_1 : P_3 \to P_3$, where for every polynomial $p(x)$ in P_3, we have $T_1(p(x)) = xp'(x)$. We have already shown that the matrix for T_1 with respect to the basis $\mathcal{B} = \{1, x, x^2, x^3\}$ of P_3 is given by

$$A_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Consider next the linear operator $T_2 : P_3 \to P_3$, where for every polynomial $q(x) = q_0 + q_1x + q_2x^2 + q_3x^3$ in P_3, we have

$$T_2(q(x)) = q(1 + x) = q_0 + q_1(1 + x) + q_2(1 + x)^2 + q_3(1 + x)^3.$$

We have $T_2(1) = 1$, $T_2(x) = 1 + x$, $T_2(x^2) = 1 + 2x + x^2$ and $T_2(x^3) = 1 + 3x + 3x^2 + x^3$, so that the matrix for T_2 with respect to $\mathcal{B}$ is given by

$$A_2 = (\,[T_2(1)]_{\mathcal{B}} \quad [T_2(x)]_{\mathcal{B}} \quad [T_2(x^2)]_{\mathcal{B}} \quad [T_2(x^3)]_{\mathcal{B}}\,) = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Consider now the composition $T = T_2 \circ T_1 : P_3 \to P_3$. Let A denote the matrix for T with respect to $\mathcal{B}$. By Proposition 8T, we have

$$A = A_2A_1 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 2 & 3 \\ 0 & 1 & 4 & 9 \\ 0 & 0 & 2 & 9 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Suppose that $p(x) = p_0 + p_1x + p_2x^2 + p_3x^3$. Then

$$[p(x)]_{\mathcal{B}} = \begin{pmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \end{pmatrix} \qquad \text{and} \qquad A[p(x)]_{\mathcal{B}} = \begin{pmatrix} 0 & 1 & 2 & 3 \\ 0 & 1 & 4 & 9 \\ 0 & 0 & 2 & 9 \\ 0 & 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \end{pmatrix} = \begin{pmatrix} p_1 + 2p_2 + 3p_3 \\ p_1 + 4p_2 + 9p_3 \\ 2p_2 + 9p_3 \\ 3p_3 \end{pmatrix},$$

so that $T(p(x)) = (p_1 + 2p_2 + 3p_3) + (p_1 + 4p_2 + 9p_3)x + (2p_2 + 9p_3)x^2 + 3p_3x^3$. We can check this directly by noting that

$$\begin{aligned} T(p(x)) = T_2(T_1(p(x))) = T_2(p_1x + 2p_2x^2 + 3p_3x^3) &= p_1(1 + x) + 2p_2(1 + x)^2 + 3p_3(1 + x)^3 \\ &= (p_1 + 2p_2 + 3p_3) + (p_1 + 4p_2 + 9p_3)x + (2p_2 + 9p_3)x^2 + 3p_3x^3. \end{aligned}$$

EXAMPLE 8.8.5. Consider the linear operator $T : \mathbb{R}^2 \to \mathbb{R}^2$, given by $T(x_1, x_2) = (2x_1 + x_2, x_1 + 3x_2)$ for every $(x_1, x_2) \in \mathbb{R}^2$. We have already shown that the matrix for T with respect to the basis $\mathcal{B} = \{(1, 0), (1, 1)\}$ of $\mathbb{R}^2$ is given by

$$A = \begin{pmatrix} 1 & -1 \\ 1 & 4 \end{pmatrix}.$$

Consider the linear operator $T^2 : \mathbb{R}^2 \to \mathbb{R}^2$. By Proposition 8T, the matrix for T^2 with respect to $\mathcal{B}$ is given by

$$A^2 = \begin{pmatrix} 1 & -1 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 4 \end{pmatrix} = \begin{pmatrix} 0 & -5 \\ 5 & 15 \end{pmatrix}.$$

Suppose that $(x_1, x_2) \in \mathbb{R}^2$. Then

$$[(x_1, x_2)]_{\mathcal{B}} = \begin{pmatrix} x_1 - x_2 \\ x_2 \end{pmatrix} \qquad \text{and} \qquad A^2[(x_1, x_2)]_{\mathcal{B}} = \begin{pmatrix} 0 & -5 \\ 5 & 15 \end{pmatrix} \begin{pmatrix} x_1 - x_2 \\ x_2 \end{pmatrix} = \begin{pmatrix} -5x_2 \\ 5x_1 + 10x_2 \end{pmatrix},$$

so that $T(x_1, x_2) = -5x_2(1, 0) + (5x_1 + 10x_2)(1, 1) = (5x_1 + 5x_2, 5x_1 + 10x^2)$. The reader is invited to check this directly.

A simple consequence of Propositions 8N and 8T is the following result concerning inverse linear transformations.

PROPOSITION 8U. *Suppose that $T : V \to V$ is a linear operator on a finite dimensional real vector space V with basis $\mathcal{B}$. Suppose further that A is the matrix for the linear operator T with respect to the basis $\mathcal{B}$. Then T is one-to-one if and only if A is invertible. Furthermore, if T is one-to-one, then A^{-1} is the matrix for the inverse linear operator $T^{-1} : V \to V$ with respect to the basis $\mathcal{B}$.*

Proof. Simply note that T is one-to-one if and only if the system $A\mathbf{x} = \mathbf{0}$ has only the trivial solution $\mathbf{x} = \mathbf{0}$. The last assertion follows easily from Proposition 8T, since if A' denotes the matrix for the inverse linear operator T^{-1} with respect to $\mathcal{B}$, then we must have $A'A = I$, the matrix for the identity operator $T^{-1} \circ T$ with respect to $\mathcal{B}$. ◯

Example 8.8.6. Consider the linear operator $T : P_3 \to P_3$, where for every $q(x) = q_0 + q_1x + q_2x^2 + q_3x^3$ in P_3, we have

$$T(q(x)) = q(1 + x) = q_0 + q_1(1 + x) + q_2(1 + x)^2 + q_3(1 + x)^3.$$

We have already shown that the matrix for T with respect to the basis $\mathcal{B} = \{1, x, x^2, x^3\}$ is given by

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

This matrix is invertible, so it follows that T is one-to-one. Furthermore, it can be checked that

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

Suppose that $p(x) = p_0 + p_1x + p_2x^2 + p_3x^3$. Then

$$[p(x)]_{\mathcal{B}} = \begin{pmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \end{pmatrix} \qquad \text{and} \qquad A^{-1}[p(x)]_{\mathcal{B}} = \begin{pmatrix} 1 & -1 & 1 & -1 \\ 0 & 1 & -2 & 3 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \end{pmatrix} = \begin{pmatrix} p_0 - p_1 + p_2 - p_3 \\ p_1 - 2p_2 + 3p_3 \\ p_2 - 3p_3 \\ p_3 \end{pmatrix},$$

so that

$$\begin{aligned} T^{-1}(p(x)) &= (p_0 - p_1 + p_2 - p_3) + (p_1 - 2p_2 + 3p_3)x + (p_2 - 3p_3)x^2 + p_3x^3 \\ &= p_0 + p_1(x - 1) + p_2(x^2 - 2x + 1) + p_3(x^3 - 3x^2 + 3x - 1) \\ &= p_0 + p_1(x - 1) + p_2(x - 1)^2 + p_3(x - 1)^3 = p(x - 1). \end{aligned}$$

8.9. Change of Basis

Suppose that V is a finite dimensional real vector space, with one basis $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ and another basis $\mathcal{B}' = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$. Suppose that $T : V \to V$ is a linear operator on V. Let A denote the matrix for T with respect to the basis $\mathcal{B}$, and let A' denote the matrix for T with respect to the basis $\mathcal{B}'$. If $\mathbf{v} \in V$ and $T(\mathbf{v}) = \mathbf{w}$, then

$$[\mathbf{w}]_{\mathcal{B}} = A[\mathbf{v}]_{\mathcal{B}} \tag{10}$$

and

$$[\mathbf{w}]_{\mathcal{B}'} = A'[\mathbf{v}]_{\mathcal{B}'}. \tag{11}$$

We wish to find the relationship between A' and A.

Recall Proposition 8J, that if

$$P = (\, [\mathbf{u}_1]_{\mathcal{B}} \quad \ldots \quad [\mathbf{u}_n]_{\mathcal{B}} \,)$$

denotes the transition matrix from the basis $\mathcal{B}'$ to the basis $\mathcal{B}$, then

$$[\mathbf{v}]_{\mathcal{B}} = P[\mathbf{v}]_{\mathcal{B}'} \qquad \text{and} \qquad [\mathbf{w}]_{\mathcal{B}} = P[\mathbf{w}]_{\mathcal{B}'}. \tag{12}$$

Note that the matrix P can also be interpreted as the matrix for the identity operator $I : V \to V$ with respect to the bases $\mathcal{B}'$ and $\mathcal{B}$. It is easy to see that the matrix P is invertible, and

$$P^{-1} = (\, [\mathbf{v}_1]_{\mathcal{B}'} \quad \ldots \quad [\mathbf{v}_n]_{\mathcal{B}'} \,)$$

denotes the transition matrix from the basis $\mathcal{B}$ to the basis $\mathcal{B}'$, and can also be interpreted as the matrix for the identity operator $I : V \to V$ with respect to the bases $\mathcal{B}$ and $\mathcal{B}'$.

Combining (10) and (12), we conclude that

$$[\mathbf{w}]_{\mathcal{B}'} = P^{-1}[\mathbf{w}]_{\mathcal{B}} = P^{-1}A[\mathbf{v}]_{\mathcal{B}} = P^{-1}AP[\mathbf{v}]_{\mathcal{B}'}.$$

Comparing this with (11), we conclude that

$$P^{-1}AP = A'. \tag{13}$$

This implies that

$$A = PA'P^{-1}. \tag{14}$$

Remark. We can use the notation

$$A = [T]_{\mathcal{B}} \qquad \text{and} \qquad A' = [T]_{\mathcal{B}'}$$

to denote that A and A' are the matrices for T with respect to the basis $\mathcal{B}$ and with respect to the basis $\mathcal{B}'$ respectively. We can also write

$$P = [I]_{\mathcal{B},\mathcal{B}'}$$

to denote that P is the transition matrix from the basis $\mathcal{B}'$ to the basis $\mathcal{B}$, so that

$$P^{-1} = [I]_{\mathcal{B}',\mathcal{B}}.$$

Then (13) and (14) become respectively

$$[I]_{\mathcal{B}',\mathcal{B}}[T]_{\mathcal{B}}[I]_{\mathcal{B},\mathcal{B}'} = [T]_{\mathcal{B}'} \qquad \text{and} \qquad [I]_{\mathcal{B},\mathcal{B}'}[T]_{\mathcal{B}'}[I]_{\mathcal{B}',\mathcal{B}} = [T]_{\mathcal{B}}.$$

We have proved the following result.

PROPOSITION 8V. *Suppose that $T : V \to V$ is a linear operator on a finite dimensional real vector space V, with bases $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ and $\mathcal{B}' = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$. Suppose further that A and A' are the matrices for T with respect to the basis $\mathcal{B}$ and with respect to the basis $\mathcal{B}'$ respectively. Then*

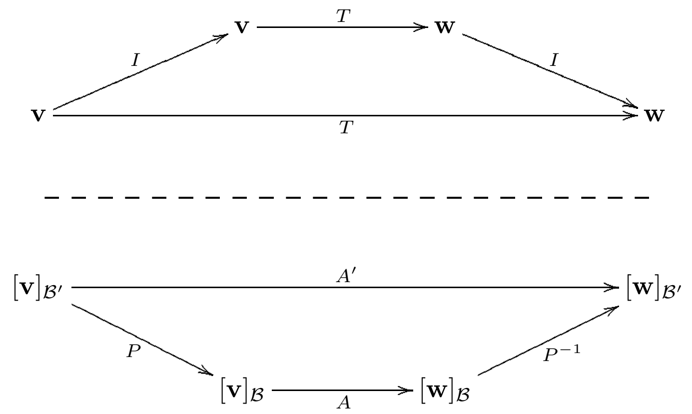
$$P^{-1}AP = A' \qquad \textit{and} \qquad A' = PAP^{-1},$$

where

$$P = (\,[\mathbf{u}_1]_{\mathcal{B}} \quad \ldots \quad [\mathbf{u}_n]_{\mathcal{B}}\,)$$

denotes the transition matrix from the basis $\mathcal{B}'$ to the basis $\mathcal{B}$.

Remarks. (1) We have the following picture.



(2) The idea can be extended to the case of linear transformations $T : V \to W$ from a finite dimensional real vector space into another, with a change of basis in V and a change of basis in W.

Example 8.9.1. Consider the vector space P_3 of all polynomials with real coefficients and degree at most 3, with bases $\mathcal{B} = \{1, x, x^2, x^3\}$ and $\mathcal{B}' = \{1, 1+x, 1+x+x^2, 1+x+x^2+x^3\}$. Consider also the linear operator $T : P_3 \to P_3$, where for every polynomial $p(x) = p_0 + p_1x + p_2x^2 + p_3x^3$, we have $T(p(x)) = (p_0+p_1) + (p_1+p_2)x + (p_2+p_3)x^2 + (p_0+p_3)x^3$. Let A denote the matrix for T with respect to the basis $\mathcal{B}$. Then $T(1) = 1 + x^3$, $T(x) = 1 + x$, $T(x^2) = x + x^2$ and $T(x^3) = x^2 + x^3$, and so

$$A = (\,[T(1)]_{\mathcal{B}} \quad [T(x)]_{\mathcal{B}} \quad [T(x^2)]_{\mathcal{B}} \quad [T(x^3)]_{\mathcal{B}}\,) = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix}.$$

Next, note that the transition matrix from the basis $\mathcal{B}'$ to the basis $\mathcal{B}$ is given by

$$P = (\,[1]_{\mathcal{B}} \quad [1+x]_{\mathcal{B}} \quad [1+x+x^2]_{\mathcal{B}} \quad [1+x+x^2+x^3]_{\mathcal{B}}\,) = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

It can be checked that

$$P^{-1} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

and so

$$A' = P^{-1}AP = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ -1 & -1 & 0 & 0 \\ 1 & 1 & 1 & 2 \end{pmatrix}$$

is the matrix for T with respect to the basis $\mathcal{B}'$. It follows that

$$\begin{aligned} T(1) &= 1 - (1 + x + x^2) + (1 + x + x^2 + x^3) = 1 + x^3, \\ T(1 + x) &= 1 + (1 + x) - (1 + x + x^2) + (1 + x + x^2 + x^3) = 2 + x + x^3, \\ T(1 + x + x^2) &= (1 + x) + (1 + x + x^2 + x^3) = 2 + 2x + x^2 + x^3, \\ T(1 + x + x^2 + x^3) &= 2(1 + x + x^2 + x^3) = 2 + 2x + 2x^2 + 2x^3. \end{aligned}$$

These can be verified directly.

8.10. Eigenvalues and Eigenvectors

Definition. Suppose that $T : V \to V$ is a linear operator on a finite dimensional real vector space V. Then any real number $\lambda \in \mathbb{R}$ is called an eigenvalue of T if there exists a non-zero vector $\mathbf{v} \in V$ such that $T(\mathbf{v}) = \lambda\mathbf{v}$. This non-zero vector $\mathbf{v} \in V$ is called an eigenvector of T corresponding to the eigenvalue λ.

The purpose of this section is to show that the problem of eigenvalues and eigenvectors of the linear operator T can be reduced to the problem of eigenvalues and eigenvectors of the matrix for T with respect to any basis $\mathcal{B}$ of V. The starting point of our argument is the following theorem, the proof of which is left as an exercise.

PROPOSITION 8W. *Suppose that $T : V \to V$ is a linear operator on a finite dimensional real vector space V, with bases $\mathcal{B}$ and $\mathcal{B}'$. Suppose further that A and A' are the matrices for T with respect to the basis $\mathcal{B}$ and with respect to the basis $\mathcal{B}'$ respectively. Then*
(a) $\det A = \det A'$;
(b) A and A' have the same rank;
(c) A and A' have the same characteristic polynomial;
(d) A and A' have the same eigenvalues; and
(e) the dimension of the eigenspace of A corresponding to an eigenvalue λ is equal to the dimension of the eigenspace of A' corresponding to λ.

We also state without proof the following result.

PROPOSITION 8X. *Suppose that $T : V \to V$ is a linear operator on a finite dimensional real vector space V. Suppose further that A is the matrix for T with respect to a basis $\mathcal{B}$ of V. Then*
(a) the eigenvalues of T are precisely the eigenvalues of A; and
(b) a vector $\mathbf{u} \in V$ is an eigenvector of T corresponding to an eigenvalue λ if and only if the coordinate matrix $[\mathbf{u}]_{\mathcal{B}}$ is an eigenvector of A corresponding to the eigenvalue λ.

Suppose now that A is the matrix for a linear operator $T : V \to V$ on a finite dimensional real vector space V with respect to a basis $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. If A can be diagonalized, then there exists an invertible matrix P such that

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

is a diagonal matrix. Furthermore, the columns of P are eigenvectors of A, and so are the coordinate matrices of eigenvectors of T with respect to the basis $\mathcal{B}$. In other words,

$$P = (\, [\mathbf{u}_1]_{\mathcal{B}} \quad \ldots \quad [\mathbf{u}_n]_{\mathcal{B}} \,),$$

where $\mathcal{B}' = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ is a basis of V consiting of eigenvectors of T. Furthermore, P is the transition matrix from the basis $\mathcal{B}'$ to the basis $\mathcal{B}$. It follows that the matrix for T with respect to the basis $\mathcal{B}'$ is given by

$$D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix},$$

where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of T.

Example 8.10.1. Consider the vector space P_2 of all polynomials with real coefficients and degree at most 2, with basis $\mathcal{B} = \{1, x, x^2\}$. Consider also the linear operator $T : P_2 \to P_2$, where for every polynomial $p(x) = p_0 + p_1x + p_2x^2$, we have $T(p(x)) = (5p_0 - 2p_1) + (6p_1 + 2p_2 - 2p_0)x + (2p_1 + 7p_2)x^2$. Then $T(1) = 5 - 2x$, $T(x) = -2 + 6x + 2x^2$ and $T(x^2) = 2x + 7x^2$, so that the matrix for T with respect to the basis $\mathcal{B}$ is given by

$$A = (\, [T(1)]_{\mathcal{B}} \quad [T(x)]_{\mathcal{B}} \quad [T(x^2)]_{\mathcal{B}} \,) = \begin{pmatrix} 5 & -2 & 0 \\ -2 & 6 & 2 \\ 0 & 2 & 7 \end{pmatrix}.$$

It is a simple exercise to show that the matrix A has eigenvalues 3, 6, 9, with corresponding eigenvectors

$$\mathbf{x}_1 = \begin{pmatrix} 2 \\ 2 \\ -1 \end{pmatrix}, \qquad \mathbf{x}_2 = \begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix}, \qquad \mathbf{x}_3 = \begin{pmatrix} -1 \\ 2 \\ 2 \end{pmatrix},$$

so that writing

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

we have

$$P^{-1}AP = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 9 \end{pmatrix}.$$

Now let $\mathcal{B}' = \{p_1(x), p_2(x), p_3(x)\}$, where

$$[p_1(x)]_{\mathcal{B}} = \begin{pmatrix} 2 \\ 2 \\ -1 \end{pmatrix}, \qquad [p_2(x)]_{\mathcal{B}} = \begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix}, \qquad [p_3(x)]_{\mathcal{B}} = \begin{pmatrix} -1 \\ 2 \\ 2 \end{pmatrix}.$$

Then P is the transition matrix from the basis $\mathcal{B}'$ to the basis $\mathcal{B}$, and D is the matrix for T with respect to the basis $\mathcal{B}'$. Clearly $p_1(x) = 2 + 2x - x^2$, $p_2(x) = 2 - x + 2x^2$ and $p_3(x) = -1 + 2x + 2x^2$. Note now that

$$\begin{aligned}
T(p_1(x)) &= T(2 + 2x - x^2) = 6 + 6x - 3x^2 = 3p_1(x), \\
T(p_2(x)) &= T(2 - x + 2x^2) = 12 - 6x + 12x^2 = 6p_2(x), \\
T(p_3(x)) &= T(-1 + 2x + 2x^2) = -9 + 18x + 18x^2 = 9p_3(x).
\end{aligned}$$

Problems for Chapter 8

1. Consider the transformation $T : \mathbb{R}^3 \to \mathbb{R}^4$, given by

$$T(x_1, x_2, x_3) = (x_1 + x_2 + x_3, x_2 + x_3, 3x_1 + x_2, 2x_2 + x_3)$$

for every $(x_1, x_2, x_3) \in \mathbb{R}^3$.
a) Find the standard matrix A for T.
b) By reducing A to row echelon form, determine the dimension of the kernel of T and the dimension of the range of T.

2. Consider a linear operator $T : \mathbb{R}^3 \to \mathbb{R}^3$ with standard matrix

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

Let $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ denote the standard basis for $\mathbb{R}^3$.
a) Find $T(\mathbf{e}_j)$ for every $j = 1, 2, 3$.
b) Find $T(2\mathbf{e}_1 + 5\mathbf{e}_2 + 3\mathbf{e}_3)$.
c) Is T invertible? Justify your assertion.

3. Consider the linear operator $T : \mathbb{R}^2 \to \mathbb{R}^2$ with standard matrix

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

a) Find the image under T of the line $x_1 + 2x_2 = 3$.
b) Find the image under T of the circle $x_1^2 + x_2^2 = 1$.

4. For each of the following, determine whether the given transformation is linear:
a) $T : V \to \mathbb{R}$, where V is a real inner product space and $T(\mathbf{u}) = \|\mathbf{u}\|$.
b) $T : \mathcal{M}_{2,2}(\mathbb{R}) \to \mathcal{M}_{2,3}(\mathbb{R})$, where $B \in \mathcal{M}_{2,3}(\mathbb{R})$ is fixed and $T(A) = AB$.
c) $T : \mathcal{M}_{3,4}(\mathbb{R}) \to \mathcal{M}_{4,3}(\mathbb{R})$, where $T(A) = A^t$.
d) $T : P_2 \to P_2$, where $T(p_0 + p_1x + p_2x^2) = p_0 + p_1(2 + x) + p_2(2 + x)^2$.
e) $T : P_2 \to P_2$, where $T(p_0 + p_1x + p_2x^2) = p_0 + p_1x + (p_2 + 1)x^2$.

5. Suppose that $T : \mathbb{R}^3 \to \mathbb{R}^3$ is a linear transformation satisfying the conditions $T(1, 0, 0) = (2, 4, 1)$, $T(1, 1, 0) = (3, 0, 2)$ and $T(1, 1, 1) = (1, 4, 6)$.
a) Evaluate $T(5, 3, 2)$.
b) Find $T(x_1, x_2, x_3)$ for every $(x_1, x_2, x_3) \in \mathbb{R}^3$.

6. Suppose that $T : \mathbb{R}^3 \to \mathbb{R}^3$ is orthogonal projection onto the x_1x_2-plane.
a) Find the standard matrix A for T.
b) Find A^2.
c) Show that $T \circ T = T$.

7. Consider the bases $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ and $\mathcal{C} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ of $\mathbb{R}^3$, where $\mathbf{u}_1 = (2, 1, 1)$, $\mathbf{u}_2 = (2, -1, 1)$, $\mathbf{u}_3 = (1, 2, 1)$, $\mathbf{v}_1 = (3, 1, -5)$, $\mathbf{v}_2 = (1, 1, -3)$ and $\mathbf{v}_3 = (-1, 0, 2)$.
a) Find the transition matrix from the basis $\mathcal{C}$ to the basis $\mathcal{B}$.
b) Find the transition matrix from the basis $\mathcal{B}$ to the basis $\mathcal{C}$.
c) Show that the matrices in parts (a) and (b) are inverses of each other.
d) Compute the coordinate matrix $[\mathbf{u}]_{\mathcal{C}}$, where $\mathbf{u} = (-5, 8, -5)$.
e) Use the transition matrix to compute the coordinate matrix $[\mathbf{u}]_{\mathcal{B}}$.
f) Compute the coordinate matrix $[\mathbf{u}]_{\mathcal{B}}$ directly and compare it to your answer in part (e).

8. Consider the bases $\mathcal{B} = \{p_1, p_2\}$ and $\mathcal{C} = \{q_1, q_2\}$ of P_1, where $p_1 = 2$, $p_2 = 3 + 2x$, $q_1 = 6 + 3x$ and $q_2 = 10 + 2x$.
   a) Find the transition matrix from the basis $\mathcal{C}$ to the basis $\mathcal{B}$.
   b) Find the transition matrix from the basis $\mathcal{B}$ to the basis $\mathcal{C}$.
   c) Show that the matrices in parts (a) and (b) are inverses of each other.
   d) Compute the coordinate matrix $[p]_{\mathcal{C}}$, where $p = -4 + x$.
   e) Use the transition matrix to compute the coordinate matrix $[p]_{\mathcal{B}}$.
   f) Compute the coordinate matrix $[p]_{\mathcal{B}}$ directly and compare it to your answer in part (e).

9. Let V be the real vector space spanned by the functions $f_1 = \sin x$ and $f_2 = \cos x$.
   a) Show that $g_1 = 2\sin x + \cos x$ and $g_2 = 3\cos x$ form a basis of V.
   b) Find the transition matrix from the basis $\mathcal{C} = \{g_1, g_2\}$ to the basis $\mathcal{B} = \{f_1, f_2\}$ of V.
   c) Compute the coordinate matrix $[f]_{\mathcal{C}}$, where $f = 2\sin x - 5\cos x$.
   d) Use the transition matrix to compute the coordinate matrix $[f]_{\mathcal{B}}$.
   e) Compute the coordinate matrix $[f]_{\mathcal{B}}$ directly and compare it to your answer in part (d).

10. Let P be the transition matrix from a basis $\mathcal{C}$ to another basis $\mathcal{B}$ of a real vector space V. Explain why P is invertible.

11. For each of the following linear transformations T, find $\ker(T)$ and $R(T)$, and verify the Rank-nullity theorem:
   a) $T : \mathbb{R}^3 \to \mathbb{R}^3$, with standard matrix $A = \begin{pmatrix} 1 & -1 & 3 \\ 5 & 6 & -4 \\ 7 & 4 & 2 \end{pmatrix}$.
   b) $T : P_3 \to P_2$, where $T(p(x)) = p'(x)$, the formal derivative.
   c) $T : P_1 \to \mathbb{R}$, where $T(p(x)) = \displaystyle\int_0^1 p(x)\,\mathrm{d}x$.

12. For each of the following, determine whether the linear operator $T : \mathbb{R}^n \to \mathbb{R}^n$ is one-to-one. If so, find also the inverse linear operator $T^{-1} : \mathbb{R}^n \to \mathbb{R}^n$:
   a) $T(x_1, x_2, x_3, \ldots, x_n) = (x_2, x_1, x_3, \ldots, x_n)$
   b) $T(x_1, x_2, x_3, \ldots, x_n) = (x_2, x_3, \ldots, x_n, x_1)$
   c) $T(x_1, x_2, x_3, \ldots, x_n) = (x_2, x_2, x_3, \ldots, x_n)$

13. Consider the operator $T : \mathbb{R}^2 \to \mathbb{R}^2$, where $T(x_1, x_2) = (x_1 + kx_2, -x_2)$ for every $(x_1, x_2) \in \mathbb{R}^2$. Here $k \in \mathbb{R}$ is fixed.
   a) Show that T is a linear operator.
   b) Show that T is one-to-one.
   c) Find the inverse linear operator $T^{-1} : \mathbb{R}^2 \to \mathbb{R}^2$.

14. Consider the linear transformation $T : P_2 \to P_1$, where $T(p_0 + p_1x + p_2x^2) = (p_0 + p_2) + (2p_0 + p_1)x$ for every polynomial $p_0 + p_1x + p_2x^2$ in P_2.
   a) Find the matrix for T with respect to the bases $\{1, x, x^2\}$ and $\{1, x\}$.
   b) Find $T(2 + 3x + 4x^2)$ by using the matrix A.
   c) Use the matrix A to recover the formula $T(p_0 + p_1x + p_2x^2) = (p_0 + p_2) + (2p_0 + p_1)x$.

15. Consider the linear operator $T : \mathbb{R}^2 \to \mathbb{R}^2$, where $T(x_1, x_2) = (x_1 - x_2, x_1 + x_2)$ for every $(x_1, x_2) \in \mathbb{R}^2$.
   a) Find the matrix A for T with respect to the basis $\{(1, 1), (-1, 0)\}$ of $\mathbb{R}^2$.
   b) Use the matrix A to recover the formula $T(x_1, x_2) = (x_1 - x_2, x_1 + x_2)$.
   c) Is T one-to-one? If so, use the matrix A to find the inverse linear operator $T^{-1} : \mathbb{R}^2 \to \mathbb{R}^2$.

16. Consider the real vector space of all real sequences $\mathbf{x} = (x_1, x_2, x_3, \ldots)$ such that the series

$$\sum_{n=1}^{\infty} x_n$$

is convergent.

a) Show that the transformation $T : V \to \mathbb{R}$, given by

$$T(\mathbf{x}) = \sum_{n=1}^{\infty} x_n$$

for every $\mathbf{x} \in V$, is a linear transformation.

b) Is the linear transformation T one-to-one? If so, give a proof. If not, find two distinct vectors $\mathbf{x}, \mathbf{y} \in V$ such that $T(\mathbf{x}) = T(\mathbf{y})$.

17. Suppose that $T_1 : \mathbb{R}^2 \to \mathbb{R}^2$ and $T_2 : \mathbb{R}^2 \to \mathbb{R}^2$ are linear operators such that

$$T_1(x_1, x_2) = (x_1 + x_2, x_1 - x_2) \qquad \text{and} \qquad T_2(x_1, x_2) = (2x_1 + x_2, x_1 - 2x_2)$$

for every $(x_1, x_2) \in \mathbb{R}^2$.

a) Show that T_1 and T_2 are one-to-one.

b) Find the formulas for T_1^{-1}, T_2^{-1} and $(T_2 \circ T_1)^{-1}$.

c) Verify that $(T_2 \circ T_1)^{-1} = T_1^{-1} \circ T_2^{-1}$.

18. Consider the transformation $T : P_1 \to \mathbb{R}^2$, where $T(p(x)) = (p(0), p(1))$ for every polynomial $p(x)$ in P_1.

a) Find $T(1 - 2x)$.

b) Show that T is a linear transformation.

c) Show that T is one-to-one.

d) Find $T^{-1}(2, 3)$, and sketch its graph.

19. Suppose that V and W are finite dimensional real vector spaces with $\dim V > \dim W$. Suppose further that $T : V \to W$ is a linear transformation. Explain why T cannot be one-to-one.

20. Suppose that

$$A = \begin{pmatrix} 1 & 3 & -1 \\ 2 & 0 & 5 \\ 6 & -2 & 4 \end{pmatrix}$$

is the matrix for a linear operator $T : P_2 \to P_2$ with respect to the basis $\mathcal{B} = \{p_1(x), p_2(x), p_3(x)\}$ of P_2, where $p_1(x) = 3x + 3x^2$, $p_2(x) = -1 + 3x + 2x^2$ and $p_3(x) = 3 + 7x + 2x^2$.

a) Find $[T(p_1(x))]_{\mathcal{B}}$, $[T(p_2(x))]_{\mathcal{B}}$ and $[T(p_3(x))]_{\mathcal{B}}$.

b) Find $T(p_1(x))$, $T(p_2(x))$ and $T(p_3(x))$.

c) Find a formula for $T(p_0 + p_1 x + p_2 x^2)$.

d) Use the formula in part (c) to compute $T(1 + x^2)$.

21. Suppose that $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ is a basis for a real vector space V. Suppose that $T : V \to V$ is a linear operator, with $T(\mathbf{v}_1) = \mathbf{v}_2$, $T(\mathbf{v}_2) = \mathbf{v}_4$, $T(\mathbf{v}_3) = \mathbf{v}_1$ and $T(\mathbf{v}_4) = \mathbf{v}_3$.

a) Find the matrix for T with respect to the basis $\mathcal{B}$.

b) Is T one-to-one? If so, describe its inverse.

22. Let P_k denote the vector space of all polynomials with real coefficients and degree at most k. Consider P_2 with basis $\mathcal{B} = \{1, x, x^2\}$ and P_3 with basis $\mathcal{C} = \{1, x, x^2, x^3\}$. We define $T_1 : P_2 \to P_3$ and $T_2 : P_3 \to P_2$ as follows. For every polynomial $p(x) = a_0 + a_1x + a_2x^2$ in P_2, we have $T_1(p(x)) = xp(x) = a_0x + a_1x^2 + a_2x^3$. For every polynomial $q(x)$ in P_3, we have $T_2(q(x)) = q'(x)$, the formal derivative of $q(x)$ with respect to the variable x.
 a) Show that $T_1 : P_2 \to P_3$ and $T_2 : P_3 \to P_2$ are linear transformations.
 b) Find $T_1(1)$, $T_1(x)$, $T_1(x^2)$, and compute the matrix A_1 for $T_1 : P_2 \to P_3$ with respect to the bases $\mathcal{B}$ and $\mathcal{C}$.
 c) Find $T_2(1)$, $T_2(x)$, $T_2(x^2)$, $T_2(x^3)$, and compute the matrix A_2 for $T_2 : P_3 \to P_2$ with respect to the bases $\mathcal{C}$ and $\mathcal{B}$.
 d) Let $T = T_2 \circ T_1$. Find $T(1)$, $T(x)$, $T(x^2)$, and compute the matrix A for $T : P_2 \to P_2$ with respect to the basis $\mathcal{B}$. Verify that $A = A_2A_1$.

23. Suppose that $T : V \to V$ is a linear operator on a real vector space V with basis $\mathcal{B}$. Suppose that for every $\mathbf{v} \in V$, we have

$$[T(\mathbf{v})]_{\mathcal{B}} = \begin{pmatrix} x_1 - x_2 + x_3 \\ x_1 + x_2 \\ x_1 - x_2 \end{pmatrix} \qquad \text{and} \qquad [\mathbf{v}]_{\mathcal{B}} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

 a) Find the matrix for T with respect to the basis $\mathcal{B}$.
 b) Is T one-to-one? If so, describe its inverse.

24. For each of the following, let V be the subspace with basis $\mathcal{B} = \{f_1(x), f_2(x), f_3(x)\}$ of the space of all real valued functions defined on $\mathbb{R}$. Let $T : V \to V$ be defined by $T(f(x)) = f'(x)$ for every function $f(x)$ in V. Find the matrix for T with respect to the basis $\mathcal{B}$:
 a) $f_1(x) = 1$, $f_2(x) = \sin x$, $f_3(x) = \cos x$
 b) $f_1(x) = \mathrm{e}^{2x}$, $f_2(x) = x\mathrm{e}^{2x}$, $f_3(x) = x^2\mathrm{e}^{2x}$

25. Let P_2 denote the vector space of all polynomials with real coefficients and degree at most 2, with basis $\mathcal{B} = \{1, x, x^2\}$. Consider the linear operator $T : P_2 \to P_2$, where for every polynomial $p(x) = a_0 + a_1x + a_2x^2$ in P_2, we have $T(p(x)) = p(2x + 1) = a_0 + a_1(2x + 1) + a_2(2x + 1)^2$.
 a) Find $T(1)$, $T(x)$, $T(x^2)$, and compute the matrix A for T with respect to the basis $\mathcal{B}$.
 b) Use the matrix A to compute $T(3 + x + 2x^2)$.
 c) Check your calculations in part (b) by computing $T(3 + x + 2x^2)$ directly.
 d) What is the matrix for $T \circ T : P_2 \to P_2$ with respect to the basis $\mathcal{B}$?
 e) Consider a new basis $\mathcal{B}' = \{1 + x, 1 + x^2, x + x^2\}$ of P_2. Using a change of basis matrix, compute the matrix for T with respect to the basis $\mathcal{B}'$.
 f) Check your answer in part (e) by computing the matrix directly.

26. Consider the linear operator $T : P_1 \to P_1$, where for every polynomial $p(x) = p_0 + p_1x$ in P_1, we have $T(p(x)) = p_0 + p_1(x + 1)$. Consider also the bases $\mathcal{B} = \{6 + 3x, 10 + 2x\}$ and $\mathcal{B}' = \{2, 3 + 2x\}$ of P_1.
 a) Find the matrix for T with respect to the basis $\mathcal{B}$.
 b) Use Proposition 8V to compute the matrix for T with respect to the basis $\mathcal{B}'$.

27. Suppose that V and W are finite dimensional real vector spaces. Suppose further that $\mathcal{B}$ and $\mathcal{B}'$ are bases for V, and that $\mathcal{C}$ and $\mathcal{C}'$ are bases for W. Show that for any linear transformation $T : V \to W$, we have

$$[I]_{\mathcal{C}',\mathcal{C}}[T]_{\mathcal{C},\mathcal{B}}[I]_{\mathcal{B},\mathcal{B}'} = [T]_{\mathcal{C}',\mathcal{B}'}.$$

28. Prove Proposition 8W.

29. Prove Proposition 8X.

30. For each of the following linear transformations $T : \mathbb{R}^3 \to \mathbb{R}^3$, find a basis $\mathcal{B}$ of $\mathbb{R}^3$ such that the matrix for T with respect to the basis $\mathcal{B}$ is a diagonal matrix:
    a) $T(x_1, x_2, x_3) = (-x_2 + x_3, -x_1 + x_3, x_1 + x_2)$
    b) $T(x_1, x_2, x_3) = (4x_1 + x_3, 2x_1 + 3x_2 + 2x_3, x_1 + 4x_3)$

31. Consider the linear operator $T : P_2 \to P_2$, where

$$T(p_0 + p_1x + p_2x^2) = (p_0 - 6p_1 + 12p_2) + (13p_1 - 30p_2)x + (9p_1 - 20p_2)x^2.$$

    a) Find the eigenvalues of T.
    b) Find a basis $\mathcal{B}$ of P_2 such that the matrix for T with respect to $\mathcal{B}$ is a diagonal matrix.