Overview

Last week introduced the important Diagonalisation Theorem:

> *An $n \times n$ matrix A is diagonalisable if and only if there is a basis for $\mathbb{R}^n$ consisting of eigenvectors of A.*

This week we'll continue our study of eigenvectors and eigenvalues, but instead of focusing just on the matrix, we'll consider the associated linear transformation.

From Lay, §5.4

Question

If we always treat a matrix as defining a linear transformation, what role does diagonalisation play?

(The version of the lecture notes posted online has more examples than will be covered in class.)

Introduction

We know that a matrix determines a linear transformation, but the converse is also true:
if $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, then T can be obtained as a matrix transformation

$$(*) \qquad T(\mathbf{x}) = A\mathbf{x} \quad \text{for all } \mathbf{x} \in \mathbb{R}^n$$

for a unique matrix A.
To construct this matrix, define

$$A = [T(\mathbf{e}_1)\ T(\mathbf{e}_2)\ \cdots\ T(\mathbf{e}_n)],$$

the $m \times n$ matrix whose columns are the images via T of the vectors of the standard basis for $\mathbb{R}^n$ (notice that $T(\mathbf{e}_i)$ is a vector in $\mathbb{R}^m$ for every $i = 1, \ldots, n$).

The matrix A is called the *standard matrix* of T.

Example 1

Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be the linear transformation defined by the formula

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x-y \\ 3x+y \\ x-y \end{bmatrix}.$$

Find the standard matrix of T.

The standard matrix of T is the matrix $[[T(\mathbf{e}_1)]_{\mathcal{E}}\ [T(\mathbf{e}_2)]_{\mathcal{E}}]$.
Since

$$T(\mathbf{e}_1) = T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}, \qquad T(\mathbf{e}_2) = T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix},$$

the standard matrix of T is the 3×2 matrix

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

Example 2

Let $A = \begin{bmatrix} 2 & 0 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. What does the linear transformation $T(\mathbf{x}) = A\mathbf{x}$ do to each of the standard basis vectors?

- The image of $\mathbf{e_1}$ is the vector $\begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} = T(\mathbf{e_1})$. Thus, we see that T multiplies any vector parallel to the x-axis by the scalar 2.
- The image of $\mathbf{e_2}$ is the vector $\begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} = T(\mathbf{e_2})$. Thus, we see that T multiplies any vector parallel to the y-axis by the scalar -1.
- The image of $\mathbf{e_3}$ is the vector $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = T(\mathbf{e_3})$. Thus, we see that T sends a vector parallel to the z-axis to a vector with equal x and z coordinates.

When we introduced the notion of coordinates, we noted that choosing different bases for our vector space gave us different coordinates. For example, suppose

$$\mathcal{E} = \{\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}\} \text{ and } \mathcal{B} = \{\mathbf{e_1}, \mathbf{e_2}, -\mathbf{e_1} + \mathbf{e_3}\}.$$

Then

$$\mathbf{e_3} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}_{\mathcal{E}} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}_{\mathcal{B}}.$$

When we say that $T\mathbf{x} = A\mathbf{x}$, we are implicitly assuming that everything is written in terms of standard $\mathcal{E}$ coordinates.

Instead, it's more precise to write

$$[T(\mathbf{x})]_{\mathcal{E}} = A[\mathbf{x}]_{\mathcal{E}} \quad \text{with} \quad A = [[T(\mathbf{e}_1)]_{\mathcal{E}} \; [T(\mathbf{e}_2)]_{\mathcal{E}} \; \cdots \; [T(\mathbf{e}_n)]_{\mathcal{E}}]$$

Every linear transformation T from $\mathbb{R}^n$ to $\mathbb{R}^m$ can be described as multiplication by its standard matrix: the standard matrix of T describes the action of T in terms of the coordinate systems on $\mathbb{R}^n$ and $\mathbb{R}^m$ given by the standard bases of these spaces.

If we start with a vector expressed in $\mathcal{E}$ coordinates, then it's convenient to represent the linear transformation T by $[T(\mathbf{x})]_{\mathcal{E}} = A[\mathbf{x}]_{\mathcal{E}}$.

However, for any sets of coordinates on the domain and codomain, we can find a matrix that represents the linear transformation in those coordinates:

$$[T(\mathbf{x})]_{\mathcal{C}} = A[\mathbf{x}]_{\mathcal{B}}$$

(Note that the domain and codomain can be described using different coordinates! This is obvious when A is an $m \times n$ matrix for $m \neq n$, but it's also true for linear transformations from $\mathbb{R}^n$ to itself.)

Example 3

For $A = \begin{bmatrix} 2 & 0 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, we saw that $[T(\mathbf{x})]_{\mathcal{E}} = A[\mathbf{x}]_{\mathcal{E}}$ acted as follows:

- T multiplies any vector parallel to the x-axis by the scalar 2.
- T multiplies any vector parallel to the y-axis by the scalar -1.
- T sends a vector parallel to the z-axis to a vector with equal x and z coordinates.

Describe the matrix B such that $[T(\mathbf{x})]_{\mathcal{B}} = A[\mathbf{x}]_{\mathcal{B}}$, where $\mathcal{B} = \{\mathbf{e_1}, \mathbf{e_2}, -\mathbf{e_1} + \mathbf{e_3}\}$.

Just as the i^{th} column of A is $[T(\mathbf{e_i})]_{\mathcal{E}}$, the i^{th} column of B will be $[T(\mathbf{b_i})]_{\mathcal{B}}$.

Since $\mathbf{e_1} = \mathbf{b_1}$, $T(\mathbf{b_1}) = 2\mathbf{b_1}$. Similarly, $T(\mathbf{b_2}) = -\mathbf{b_2}$.

Thus we see that $B = \begin{bmatrix} 2 & 0 & * \\ 0 & -1 & * \\ 0 & 0 & * \end{bmatrix}$.

The third column is the interesting one. Again, recall $\mathcal{B} = \{\mathbf{e_1}, \mathbf{e_2}, -\mathbf{e_1} + \mathbf{e_3}\}$, and

- T multiplies any vector parallel to the x-axis by the scalar 2.
- T multiplies any vector parallel to the y-axis by the scalar -1.
- T sends a vector parallel to the z-axis to a vector with equal x and z coordinates.

The 3^{rd} column of B will be $[T(\mathbf{b_3})]_{\mathcal{B}}$.

$$T(\mathbf{b_3}) = T(-\mathbf{e_1}+\mathbf{e_3}) = -T(\mathbf{e_1})+T(\mathbf{e_3}) = -2\mathbf{e_1}+(\mathbf{e_1}+\mathbf{e_3}) = -\mathbf{e_1}+\mathbf{e_3} = \mathbf{b_3}.$$

Thus we see that $B = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

Notice that in B coordinates, the matrix representing T is diagonal!

Every linear transformation $T : V \to W$ between finite dimensional vector spaces can be represented by a matrix, but the matrix representation of a linear transformation depends on the choice of bases for V and W (thus it is not unique).

This allows us to reduce many linear algebra problems concerning abstract vector spaces to linear algebra problems concerning the familiar vector spaces $\mathbb{R}^n$. This is important even for linear transformations $T : \mathbb{R}^n \to \mathbb{R}^m$ since certain choices of bases for $\mathbb{R}^n$ and $\mathbb{R}^m$ can make important properties of T more evident: to solve certain problems easily, it is important to choose the *right* coordinates.

Matrices and linear transformations

Let $T : V \to W$ be a linear transformation that maps from V to W, and suppose that we've fixed a basis $\mathcal{B} = \{\mathbf{b_1}, \ldots, \mathbf{b_n}\}$ for V and a basis $\mathcal{C} = \{\mathbf{c_1}, \ldots, \mathbf{c_m}\}$ for W.

For any vector $\mathbf{x} \in V$, the coordinate vector $[\mathbf{x}]_\mathcal{B}$ is in $\mathbb{R}^n$ and the coordinate vector of its image $[T(\mathbf{x})]_\mathcal{C}$ is in $\mathbb{R}^m$.

We want to associate a matrix M with T with the property that $M[\mathbf{x}]_\mathcal{B} = [T(\mathbf{x})]_\mathcal{C}$.

It can be helpful to organise this information with a diagram

$$\begin{tikzcd} V \ni \mathbf{x} \arrow[r, "T"] \arrow[d] & T(\mathbf{x}) \in W \arrow[d] \\ \mathbb{R}^n \ni [\mathbf{x}]_\mathcal{B} \arrow[r, "\text{multiplication by } M"'] & [T(\mathbf{x})]_\mathcal{C} \in \mathbb{R}^m \end{tikzcd}$$

where the vertical arrows represent the coordinate mappings.

Here's an example to illustrate how we might find such a matrix M:
Let $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ and $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$ be bases for two vector spaces V and W, respectively.
Let $T : V \to W$ be the linear transformation defined by

$$T(\mathbf{b}_1) = 2\mathbf{c}_1 - 3\mathbf{c}_2, \; T(\mathbf{b}_2) = -4\mathbf{c}_1 + 5\mathbf{c}_2\,.$$

Why does this define the entire linear transformation? For an arbitrary

vector $\mathbf{v} = x_1\mathbf{b}_1 + x_2\mathbf{b}_2$ in V, we define its image under T as

$$T(\mathbf{v}) = x_1 T(\mathbf{b}_1) + x_2 T(\mathbf{b}_2).$$

For example, if $\mathbf{x}$ is the vector in V given by $\mathbf{x} = 3\mathbf{b}_1 + 2\mathbf{b}_2$, so that $[\mathbf{x}]_\mathcal{B} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$, we have

$$\begin{aligned} T(\mathbf{x}) &= T(3\mathbf{b}_1 + 2\mathbf{b}_2) \\ &= 3T(\mathbf{b}_1) + 2T(\mathbf{b}_2) \\ &= 3(2\mathbf{c}_1 - 3\mathbf{c}_2) + 2(-4\mathbf{c}_1 + 5\mathbf{c}_2) \\ &= -2\mathbf{c}_1 + \mathbf{c}_2. \end{aligned}$$

Equivalently, we have

$$\begin{aligned}
[T(\mathbf{x})]_{\mathcal{C}} &= [3T(\mathbf{b}_1) + 2T(\mathbf{b}_2)]_{\mathcal{C}} \\
&= 3[T(\mathbf{b}_1)]_{\mathcal{C}} + 2[T(\mathbf{b}_2)]_{\mathcal{C}} \\
&= \begin{bmatrix} [T(\mathbf{b}_1)]_{\mathcal{C}} & [T(\mathbf{b}_2)]_{\mathcal{C}} \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix} \\
&= \begin{bmatrix} [T(\mathbf{b}_1)]_{\mathcal{C}} & [T(\mathbf{b}_2)]_{\mathcal{C}} \end{bmatrix} [\mathbf{x}]_{\mathcal{B}}
\end{aligned}$$

In this case, since $T(\mathbf{b}_1) = 2\mathbf{c}_1 - 3\mathbf{c}_2$ and $T(\mathbf{b}_2) = -4\mathbf{c}_1 + 5\mathbf{c}_2$ we have

$$[T(\mathbf{b}_1)]_{\mathcal{C}} = \begin{bmatrix} 2 \\ -3 \end{bmatrix} \quad \text{and} \quad [T(\mathbf{b}_2)]_{\mathcal{C}} = \begin{bmatrix} -4 \\ 5 \end{bmatrix}$$

and so

$$\begin{aligned}
[T(\mathbf{x})]_{\mathcal{C}} &= \begin{bmatrix} 2 & -4 \\ -3 & 5 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix} \\
&= \begin{bmatrix} -2 \\ 1 \end{bmatrix}.
\end{aligned}$$

In the last page, we are not so much interested in the actual calculation but in the equation

$$[T(\mathbf{x})]_{\mathcal{C}} = \begin{bmatrix} [T(\mathbf{b}_1)]_{\mathcal{C}} & [T(\mathbf{b}_2)]_{\mathcal{C}} \end{bmatrix} [\mathbf{x}]_{\mathcal{B}}$$

This gives us the matrix M:

$$M = \begin{bmatrix} [T(\mathbf{b}_1)]_{\mathcal{C}} & [T(\mathbf{b}_2)]_{\mathcal{C}} \end{bmatrix}$$

whose columns consist of the coordinate vectors of $T(\mathbf{b}_1)$ and $T(\mathbf{b}_2)$ with respect to the basis $\mathcal{C}$ in W.

In general, when T is a linear transformation that maps from V to W where $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ is a basis for V and $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_m\}$ is a basis for W the matrix associated to T with respect to these bases is

$$M = \begin{bmatrix} [T(\mathbf{b}_1)]_{\mathcal{C}} & \cdots & [T(\mathbf{b}_n)]_{\mathcal{C}} \end{bmatrix}.$$

We write ${}_{\mathcal{C}\leftarrow\mathcal{B}}T$ for M, so that ${}_{\mathcal{C}\leftarrow\mathcal{B}}T$ has the property

$$\begin{aligned}
[T(\mathbf{x})]_{\mathcal{C}} &= \begin{bmatrix} [T(\mathbf{b}_1)]_{\mathcal{C}} & \cdots & [T(\mathbf{b}_n)]_{\mathcal{C}} \end{bmatrix} [\mathbf{x}]_{\mathcal{B}} \\
&= {}_{\mathcal{C}\leftarrow\mathcal{B}}T[\mathbf{x}]_{\mathcal{B}}.
\end{aligned}$$

The matrix ${}_{\mathcal{C}\leftarrow\mathcal{B}}T$ describes how the linear transformation T operates in terms of the coordinate systems on V and W associated to the basis $\mathcal{B}$ and $\mathcal{C}$ respectively.

NB. $\underset{\mathcal{C}\leftarrow\mathcal{B}}{T}$ is *the matrix for* T *relative to* $\mathcal{B}$ *and* $\mathcal{C}$. It depends on the choice of *both* the bases $\mathcal{B}, \mathcal{C}$. The order of $\mathcal{B}$, $\mathcal{C}$ is important.

In the case that $T : V \to V$ and $\mathcal{B} = \mathcal{C}$, $\underset{\mathcal{B}\leftarrow\mathcal{B}}{T}$ is written $[T]_\mathcal{B}$ and is the *matrix for* T *relative to* $\mathcal{B}$, or more shortly the $\mathcal{B}$-matrix of T.

So by taking bases in each space, and writing vectors with respect to these bases, T can be studied by studying the matrix associated to T with respect to these bases.

Algorithm for finding the matrix $\underset{\mathcal{C}\leftarrow\mathcal{B}}{T}$

To find the matrix $\underset{\mathcal{C}\leftarrow\mathcal{B}}{T}$ where $T : V \to W$ relative to

a basis $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ of V

a basis $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_m\}$ of W

- Find $T(\mathbf{b}_1), T(\mathbf{b}_2), \dots, T(\mathbf{b}_n)$.
- Find the coordinate vector $[T(\mathbf{b}_1)]_\mathcal{C}$ of $T(\mathbf{b}_1)$ with respect to the basis $\mathcal{C}$. This is a column vector in $\mathbb{R}^m$.
- Do this for each $T(\mathbf{b}_i)$.
- Make a matrix from these column vectors. This matrix is $\underset{\mathcal{C}\leftarrow\mathcal{B}}{T}$.

N.B. The coordinate vectors $[T(\mathbf{b}_1)]_\mathcal{C}, [T(\mathbf{b}_2)]_\mathcal{C}, \dots, [T(\mathbf{b}_n)]_\mathcal{C}$ have to be written as columns (not rows!).

Examples

Example 4

Let $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ and $\mathcal{D} = \{\mathbf{d}_1, \mathbf{d}_2\}$ be bases for vector spaces V and W respectively. $T : V \to W$ is the linear transformation with the property that

$$\begin{aligned} T(\mathbf{b}_1) &= 3\mathbf{d}_1 - 5\mathbf{d}_2, \\ T(\mathbf{b}_2) &= -\mathbf{d}_1 + 6\mathbf{d}_2, \\ T(\mathbf{b}_3) &= 4\mathbf{d}_2 \end{aligned}$$

We find the matrix $\underset{\mathcal{D}\leftarrow\mathcal{B}}{T}$ of T relative to $\mathcal{B}$ and $\mathcal{D}$.

We have

$$[T(\mathbf{b}_1)]_{\mathcal{D}} = \begin{bmatrix} 3 \\ -5 \end{bmatrix}, [T(\mathbf{b}_2)]_{\mathcal{D}} = \begin{bmatrix} -1 \\ 6 \end{bmatrix}$$

and

$$[T(\mathbf{b}_3)]_{\mathcal{D}} = \begin{bmatrix} 0 \\ 4 \end{bmatrix}$$

This gives

$$\begin{aligned} \underset{\mathcal{D}\leftarrow\mathcal{B}}{T} &= \begin{bmatrix} [T(\mathbf{b}_1)]_{\mathcal{D}} & [T(\mathbf{b}_3)]_{\mathcal{D}} & [T(\mathbf{b}_3)]_{\mathcal{D}} \end{bmatrix} \\ &= \begin{bmatrix} 3 & -1 & 0 \\ -5 & 6 & 4 \end{bmatrix}. \end{aligned}$$

Example 5

Define $T : \mathbb{P}_2 \to \mathbb{R}^2$ by

$$T(p(t)) = \begin{bmatrix} p(0) + p(1) \\ p(-1) \end{bmatrix}.$$

(a) Show that T is a linear transformation.

(b) Find the matrix $\underset{\mathcal{E}\leftarrow\mathcal{B}}{T}$ of T relative to the standard bases $\mathcal{B} = \{1, t, t^2\}$ and $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2\}$ of $\mathbb{P}_2$ and $\mathbb{R}^2$.

(a) This is an exercise for you.

combinations of the vectors in $\mathcal{E}$).

$$T(1) = \begin{bmatrix} 1+1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} = 2\mathbf{e}_1 + \mathbf{e}_2$$

$$T(t) = \begin{bmatrix} 0+1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \mathbf{e}_1 - \mathbf{e}_2$$

$$T(t^2) = \begin{bmatrix} 0+1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \mathbf{e}_1 + \mathbf{e}_2.$$

• STEP 2 We find the coordinate vectors of $T(1)$, $T(t)$, $T(t^2)$ in the basis $\mathcal{E}$:

$$[T(1)]_{\mathcal{E}} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \qquad [T(t)]_{\mathcal{E}} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \qquad [T(t^2)]_{\mathcal{E}} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

• STEP 3 We form the matrix whose columns are the coordinate vectors in step 2

$$\underset{\mathcal{E}\leftarrow\mathcal{B}}{T} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix}$$

Example 6

Let $V = \mathrm{Span}\{\sin t, \cos t\}$, and $D : V \to V$ the linear transformation $D : f \mapsto f'$. Let $\mathbf{b}_1 = \sin t, \mathbf{b}_2 = \cos t$, $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$, a basis for V. We find the matrix of T with respect to the basis $\mathcal{B}$.

- STEP 1 We have

$$D(\mathbf{b}_1) = \cos t = 0\mathbf{b}_1 + 1\mathbf{b}_2,$$
$$D(\mathbf{b}_2) = -\sin t = -1\mathbf{b}_1 + 0\mathbf{b}_2.$$

- STEP 2 From this we have

$$[D(\mathbf{b}_1)]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, [D(\mathbf{b}_2)]_{\mathcal{B}} = \begin{bmatrix} -1 \\ 0 \end{bmatrix},$$

- STEP 3 So that

$$[D]_{\mathcal{B}} = \begin{bmatrix} [T(\mathbf{b}_1)_{\mathcal{B}} & [T(\mathbf{b}_2)]_{\mathcal{B}} \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Let $f(t) = 4\sin t - 6\cos t$. We can use the matrix we have just found to get the derivative of $f(t)$. Now $[f(t)]_{\mathcal{B}} = \begin{bmatrix} 4 \\ -6 \end{bmatrix}$. Then

$$\begin{aligned} [D(f(t))]_{\mathcal{B}} &= [D]_{\mathcal{B}}[f(t)]_{\mathcal{B}} \\ &= \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 4 \\ -6 \end{bmatrix} \\ &= \begin{bmatrix} 6 \\ 4 \end{bmatrix}. \end{aligned}$$

This, of course gives

$$f'(t) = 6\sin t + 4\cos t$$

which is what we would expect.

Example 7

Let $M_{2\times 2}$ be the vector space of 2×2 matrixes and let $\mathbb{P}_2$ be the vector space of polynomials of degree at most 2. Let $T : M_{2\times 2} \to \mathbb{P}_2$ be the linear transformation given by

$$T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = a + b + c + (a - c)x + (a + d)x^2.$$

We find the matrix of T with respect to the basis $\mathcal{B} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$ for $M_{2\times 2}$ and the standard basis $\mathcal{C} = \{1, x, x^2\}$ for $\mathbb{P}_2$.

- STEP 1 We find the effect of T on each of the basis elements:

$$\begin{aligned}
T\left(\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\right) &= 1+x+x^2,\\
T\left(\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}\right) &= 1,\\
T\left(\begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix}\right) &= 1-x,\\
T\left(\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}\right) &= x^2.
\end{aligned}$$

- STEP 2 The corresponding coordinate vectors are

$$\begin{aligned}
\left[T\left(\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\right)\right]_{\mathcal{C}} &= \begin{bmatrix}1\\1\\1\end{bmatrix},\\
\left[T\left(\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}\right)\right]_{\mathcal{C}} &= \begin{bmatrix}1\\0\\0\end{bmatrix},\\
\left[T\left(\begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix}\right)\right]_{\mathcal{C}} &= \begin{bmatrix}1\\-1\\0\end{bmatrix},\\
\left[T\left(\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}\right)\right]_{\mathcal{C}} &= \begin{bmatrix}0\\0\\1\end{bmatrix}.
\end{aligned}$$

- STEP 3 Hence the matrix for T relative to the bases $\mathcal{B}$ and $\mathcal{C}$ is

$$\begin{bmatrix}1 & 1 & 1 & 0\\ 1 & 0 & -1 & 0\\ 1 & 0 & 0 & 1\end{bmatrix}.$$

Example 8

We consider the linear transformation

$$H : \mathbb{P}_2 \to M_{2\times 2}$$

given by

$$H(a + bx + cx^2) = \begin{bmatrix} a+b & a-b \\ c & c-a \end{bmatrix}$$

We find the matrix of H with respect to the standard basis $\mathcal{C} = \{1, x, x^2\}$ for $\mathbb{P}_2$ and
$\mathcal{B} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$ for $M_{2\times 2}$.

- STEP 1 We find the effect of H on each of the basis elements:

$$H(1) = \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}, \quad H(x) = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, H(x^2) = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}.$$

- STEP 2 The corresponding coordinate vectors are

$$[H(1)]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ -1 \end{bmatrix}, \quad [H(x)]_{\mathcal{B}} = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, [H(x^2)]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

- STEP 3 Hence the matrix for H relative to the bases $\mathcal{C}$ and $\mathcal{B}$ is

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}.$$

Linear transformations from V to V

The most common case is when $T : V \to V$ and $\mathcal{B} = \mathcal{C}$. In this case $\underset{\mathcal{B}\leftarrow\mathcal{B}}{T}$ is written $[T]_\mathcal{B}$ and is the *matrix for T relative to $\mathcal{B}$* or simply the *$\mathcal{B}$-matrix of T*.
The $\mathcal{B}$-matrix for $T : V \to V$ satisfies

$$[T(\mathbf{x})]_\mathcal{B} = [T]_\mathcal{B}[\mathbf{x}]_\mathcal{B}, \quad \text{for all } \mathbf{x} \in V. \tag{1}$$

$$\begin{array}{ccc} \mathbf{x} & \xrightarrow{\quad T \quad} & T(\mathbf{x}) \\ \downarrow & & \downarrow \\ [\mathbf{x}]_\mathcal{B} & \xrightarrow{\text{multiplication by } [T]_\mathcal{B}} & [T(\mathbf{x})]_\mathcal{B} \end{array}$$

Examples

Example 9

Let $T : \mathbb{P}_2 \to \mathbb{P}_2$ be the linear transformation defined by

$$T(p(x)) = p(2x - 1).$$

We find the matrix of T with respect to $\mathcal{E} = \{1, x, x^2\}$

- STEP 1 It is clear that

$$T(1) = 1, \quad T(x) = 2x - 1,$$
$$T(x^2) = (2x - 1)^2 = 1 - 4x + 4x^2$$

- STEP 2 So the coordinate vectors are

$$[T(1)]_\mathcal{E} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, [T(x)]_\mathcal{E} = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}, \left[T(x^2)\right]_\mathcal{E} = \begin{bmatrix} 1 \\ -4 \\ 4 \end{bmatrix}.$$

- STEP 3 Therefore

$$[T]_\mathcal{E} = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & -4 \\ 0 & 0 & 4 \end{bmatrix}$$

Example 10

We compute $T(3 + 2x - x^2)$ using part (a).

The coordinate vector of $p(x) = (3 + 2x - x^2)$ with respect to $\mathcal{E}$ is given by

$$[p(x)]_\mathcal{E} = \begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix}.$$

We use the relationship

$$[T(p(x))]_\mathcal{E} = [T]_\mathcal{E}[p(x)]_\mathcal{E}.$$

This gives

$$\begin{aligned}[T(3+2x-x^2)]_{\mathcal{E}} &= [T(p(x))]_{\mathcal{E}} \\ &= [T]_{\mathcal{E}}[p(x)]_{\mathcal{E}} \\ &= \begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & -4 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ 8 \\ -4 \end{bmatrix}\end{aligned}$$

It follows that $T(3+2x-x^2) = 8x - 4x^2$.

Example 11

Consider the linear transformation $F : M_{2\times 2} \to M_{2\times 2}$ given by

$$F(A) = A + A^T$$

where $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

We use the basis

$\mathcal{B} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$ for $M_{2\times 2}$ to find a matrix representation for T.

More explicitly F is given by

$$F\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} a & c \\ b & d \end{bmatrix} = \begin{bmatrix} 2a & b+c \\ b+c & 2d \end{bmatrix}$$

- <u>STEP 1</u> We find the effect of F on each of the basis elements:

$$F\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right) = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}, F\left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

$$F\left(\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\right) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, F\left(\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right) = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}.$$

- STEP 2 The corresponding coordinate vectors are

$$\left[F\left(\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}\right)\right]_{\mathcal{B}} = \begin{bmatrix}2\\0\\0\\0\end{bmatrix}, \left[F\left(\begin{bmatrix}0 & 1\\0 & 0\end{bmatrix}\right)\right]_{\mathcal{B}} = \begin{bmatrix}0\\1\\1\\0\end{bmatrix},$$

$$\left[F\left(\begin{bmatrix}0 & 0\\1 & 0\end{bmatrix}\right)\right]_{\mathcal{B}} = \begin{bmatrix}0\\1\\1\\0\end{bmatrix}, \left[F\left(\begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}\right)\right]_{\mathcal{B}} = \begin{bmatrix}0\\0\\0\\2\end{bmatrix}.$$

- STEP 3 Hence the matrix representing F is

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

Example 12

Let $V = \text{Span}\,\{e^{2x}, e^{2x}\cos x, e^{2x}\sin x\}$.
We find the matrix of the differential operator D with respect to

$$\mathcal{B} = \{e^{2x}, e^{2x}\cos x, e^{2x}\sin x\}.$$

- STEP 1 We see that

$$\begin{aligned} D(e^{2x}) &= 2e^{2x}\\ D(e^{2x}\cos x) &= 2e^{2x}\cos x - e^{2x}\sin x\\ D(e^{2x}\sin x) &= 2e^{2x}\sin x + e^{2x}\cos x \end{aligned}$$

- <u>STEP 2</u> So the coordinate vectors are

$$[D(e^{2x})]_{\mathcal{B}} = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}, [D(e^{2x}\cos x)]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix},$$

$$\text{and} \qquad [D(e^{2x}\sin x)]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}.$$

- <u>STEP 3</u> Hence

$$[D]_{\mathcal{B}} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & -1 & 2 \end{bmatrix}.$$

Example 13

We use this result to find the derivative of
$f(x) = 3e^{2x} - e^{2x}\cos x + 2e^{2x}\sin x$.

The coordinate vector of $f(x)$ is given by

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

We do this calculation using

$$[D(f)]_{\mathcal{B}} = [D]_{\mathcal{B}}[f]_{\mathcal{B}}.$$

This gives

$$\begin{aligned} [D(f)]_{\mathcal{B}} &= [D]_{\mathcal{B}}[f]_{\mathcal{B}} \\ &= \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} 6 \\ 0 \\ 5 \end{bmatrix}. \end{aligned}$$

This indicates that

$$f'(x) = 6e^{2x} + 5e^{2x}\sin x.$$

You should check this result by differentiation.

Example 14

We use the previous result to find $\int (4e^{2x} - 3e^{2x}\sin x)\, dx$

We recall that with the basis $\mathcal{B} = \{e^{2x}, e^{2x}\cos x, e^{2x}\sin x\}$ the matrix representation of the differential operator D is given by

$$[D]_{\mathcal{B}} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & -1 & 2 \end{bmatrix}.$$

We also notice that $[D]_{\mathcal{B}}$ is invertible with inverse:

$$[D]_{\mathcal{B}}^{-1} = \begin{bmatrix} 1/2 & 0 & 0 \\ 0 & 2/5 & -1/5 \\ 0 & 1/5 & 2/5 \end{bmatrix}.$$

The coordinate vector of $4e^{2x} - 3e^{2x}\sin x$ with respect to the basis $\mathcal{B}$ is given by $\begin{bmatrix} 4 \\ 0 \\ -3 \end{bmatrix}$. We use this together with the inverse of $[D]_{\mathcal{B}}$ to find the antiderivative $\int (4e^{2x} - 3e^{2x}\sin x)\, dx$:

$$[D]_{\mathcal{B}}^{-1}[4e^{2x} - 3e^{2x}]_{\mathcal{B}} = \begin{bmatrix} 1/2 & 0 & 0 \\ 0 & 2/5 & -1/5 \\ 0 & 1/5 & 2/5 \end{bmatrix} \begin{bmatrix} 4 \\ 0 \\ -3 \end{bmatrix} = \begin{bmatrix} 2 \\ 3/5 \\ -6/5 \end{bmatrix}.$$

So the antiderivative of $4e^{2x} - 3e^{2x}$ in the vector space V is $2e^{2x} + \frac{3}{5}e^{2x}\cos x - \frac{6}{5}e^{2x}\sin x$, and we can deduce that $\int (4e^{2x} - 3e^{2x}\sin x)\, dx = 2e^{2x} + \frac{3}{5}e^{2x}\cos x - \frac{6}{5}e^{2x}\sin x + C$ where C denotes a constant.

Linear transformations and diagonalisation

In an applied problem involving $\mathbb{R}^n$, a linear transformation T usually appears as a matrix transformation $\mathbf{x} \mapsto A\mathbf{x}$. If A is diagonalisable, then there is a basis $\mathcal{B}$ for $\mathbb{R}^n$ consisting of eigenvectors of A. In this case the $\mathcal{B}$-matrix for T is diagonal, and diagonalising A amounts to finding a diagonal matrix representation of $\mathbf{x} \mapsto A\mathbf{x}$.

Theorem

Suppose $A = PDP^{-1}$, where D is a diagonal $n \times n$ matrix. If $\mathcal{B}$ is the basis for $\mathbb{R}^n$ formed by the columns of P, then D is the $\mathcal{B}$-matrix for the transformation $\mathbf{x} \mapsto A\mathbf{x}$.

Proof.

Denote the columns of P by $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n$, so that $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ and

$$P = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_n \end{bmatrix}.$$

In this case, P is the change of coordinates matrix $P_{\mathcal{B}}$ where

$$P[\mathbf{x}]_{\mathcal{B}} = \mathbf{x} \quad \text{and} \quad [\mathbf{x}]_{\mathcal{B}} = P^{-1}\mathbf{x}.$$

If T is defined by $T(\mathbf{x}) = A\mathbf{x}$ for $\mathbf{x}$ in $\mathbb{R}^n$, then

$$\begin{aligned}
[T]_{\mathcal{B}} &= \begin{bmatrix} [T(\mathbf{b}_1)]_{\mathcal{B}} & \cdots & T(\mathbf{b}_n)]_{\mathcal{B}} \end{bmatrix} \\
&= \begin{bmatrix} [A\mathbf{b}_1]_{\mathcal{B}} & \cdots & [A\mathbf{b}_n]_{\mathcal{B}} \end{bmatrix} \\
&= \begin{bmatrix} P^{-1}A\mathbf{b}_1 & \cdots & P^{-1}A\mathbf{b}_n \end{bmatrix} \\
&= P^{-1}A \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_n \end{bmatrix} \\
&= P^{-1}AP = D
\end{aligned}$$

□

In the proof of the previous theorem the fact that D is diagonal is never used. In fact the following more general result holds:

If an $n \times n$ matrix A is similar to a matrix C with $A = PCP^{-1}$, then C is the $\mathcal{B}$-matrix of the transformation $\mathbf{x} \to A\mathbf{x}$ where $\mathcal{B}$ is the basis of $\mathbb{R}^n$ formed by the columns of P.

Example

Example 15

Consider the matrix $A = \begin{bmatrix} 4 & -2 \\ -1 & 3 \end{bmatrix}$. T is the linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T(\mathbf{x}) = A\mathbf{x}$. We find a basis $\mathcal{B}$ for $\mathbb{R}^2$ with the property that $[T]_{\mathcal{B}}$ is diagonal.

The first step is to find the eigenvalues and corresponding eigenspaces for A:

$$\begin{aligned}
\det(A - \lambda I) &= \det \begin{bmatrix} 4-\lambda & -2 \\ -1 & 3-\lambda \end{bmatrix} \\
&= (4-\lambda)(3-\lambda) - 2 \\
&= \lambda^2 - 7\lambda + 10 \\
&= (\lambda - 2)(\lambda - 5).
\end{aligned}$$

The eigenvalues of A are $\lambda = 2$ and $\lambda = 5$. We need to find a basis vector for each of these eigenspaces.

$$\begin{aligned} E_2 &= \operatorname{Nul}\begin{bmatrix} 2 & -2 \\ -1 & 1 \end{bmatrix} \\ &= \operatorname{Span}\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right\} \\ E_5 &= \operatorname{Nul}\begin{bmatrix} -1 & -2 \\ -1 & -2 \end{bmatrix} \\ &= \operatorname{Span}\left\{\begin{bmatrix} -2 \\ 1 \end{bmatrix}\right\} \end{aligned}$$

Put $\mathcal{B} = \left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \end{bmatrix}\right\}$.
Then $[T]_{\mathcal{B}} = D = \begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix}$, and with $P = \begin{bmatrix} 1 & -2 \\ 1 & 1 \end{bmatrix}$ and $P^{-1}AP = D$, or equivalently, $A = PDP^{-1}$.