4 The matrix of a linear transformation

4.1 From an abstract vector space to the space of column vectors

Throughout this section, we are thinking about a vector space V of dimension n over a field $\mathbb{F}$. Let $\mathcal{B} = \{b_1, \ldots, b_n\}$ be a *basis* of V over $\mathbb{F}$. This means that $b_1, \ldots, b_n$ are elements of V with the property that *every* element of V has a *unique expression* as a $\mathbb{F}$-linear combination of the elements of $\mathcal{B}$. Given $v \in V$, there exist unique elements $a_1, \ldots, a_n$ in $\mathbb{F}$ for which

$$v = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n.$$

We say that $a_1, \ldots, a_n$ are the $\mathcal{B}$-*coordinates* of v. We can use $\mathcal{B}$ to identify V with the space $\mathbb{F}^n$ of all column vectors of length n with entries from $\mathbb{F}$, via the correspondence

$$v \leftrightarrow \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

When necessary we will right this column vector as $[v]_{\mathcal{B}}$ (but we will try to avoid this cumbersome notation when we can).

Example 4.1. *Let V be the vector space of polynomials of degree at most 3 over $\mathbb{Q}$, and let $\mathcal{B} = \{x^3 + 1, x^2 + 1, x + 1, 1\}$. Then $\mathcal{B}$ is a basis of V over $\mathbb{Q}$ with respect to which the polynomial $2x^3 - 2x^2 + 3x - 2$ is represented by the column vector*

$$\begin{pmatrix} 2 \\ -2 \\ 3 \\ -5 \end{pmatrix}.$$

The general correspondence between V and $\mathbb{Q}^4$ determined by this basis is given by

$$ax^3 + bx^2 + cx + d \leftrightarrow \begin{pmatrix} a \\ b \\ c \\ d - (a + b + c) \end{pmatrix}.$$

Thus all vector spaces of dimension n over $\mathbb{F}$ can be identified with $\mathbb{F}^n$ (in multiple ways) and hence with each other. Moreover, you can check that the correspondence between V and $\mathbb{F}^n$ defined above is a linear transformation; it respects addition and multiplication by scalars. So all vector spaces of dimension n over $\mathbb{F}$ are isomorphic to each other, which means that in terms of abstract structure there is only one vector space of a given dimension over a given field. This statement should be interpreted with sensitivity to the context - it is not saying (for example) that the space of rational polynomials of degree 3 is the same thing as the space of 2×2 matrices over $\mathbb{Q}$, but that these two objects are isomorphic *as vector spaces* (not as anything else, for example one of them is a ring and the other is not).

Another thing to note is that the identification of an abstract vector space with a full space of column vectors is not canonical, it depends on a choice of basis. In the example above there is nothing about the particular column $\begin{pmatrix} 2 \\ -2 \\ 3 \\ -5 \end{pmatrix}$ that *intrinsically* connects it to the polynomial $2x^3 - 2x^2 + 3x - 2$, this connection depends on the choice of basis. A different choice of basis would connect the same polynomial to a different column vector.

4.2 The matrix of a linear transformation

Now let $T : V \to V$ be a linear transformation of V, where $\mathcal{B} = \{b_1, \ldots, b_n\}$ is a basis of V. With respect to $\mathcal{B}$, the vectors $v, b_1, \ldots, b_n$ are all represented by column vectors, and so also are their images under T. In particular if

$$v = a_1 b_1 + \cdots + a_n b_n$$

then

$$T(v) = a_1 T(b_1) + \cdots + a_n T(b_n),$$

and

$$[T(v)]_{\mathcal{B}} = a_1 [T(b_1)]_{\mathcal{B}} + \cdots + a_n [T(b_n)]_{\mathcal{B}}.$$

From the content of Lectures 2 and 3 we can observe that this is nothing but the matrix-vector product $A_{\mathcal{B}}[v]_{\mathcal{B}}$, where $A_{\mathcal{B}}$ is the $n \times n$ matrix whose columns are $[T(b_1)]_{\mathcal{B}}, \ldots, [T(b_n)]_{\mathcal{B}}$.

The key point is: $A_{\mathcal{B}}$ is called *the matrix of* T *with respect to the basis* $\mathcal{B}$, evaluating $T(v)$ for $v \in V$ is just evaluating the matrix-vector product $A_{\mathcal{B}}[v]_{\mathcal{B}}$, after using the basis $\mathcal{B}$ to move from the abstract vector space V into $\mathbb{F}^n$. The interpretation of matrix-vector multiplication discussed in Lectures 2 and 3, namely taking the linear combination of columns of the matrix determined by the entries of the vector, is particularly appropriate for discussing the matrix of a linear transformation.

Exercise 4.2. *As in the previous example let* V *be the space of rational polynomials of degree at most 4, with basis* $\{x^3 + 1, x^2 + 1, x + 1, 1\}$. *Let* $T : V \to V$ *be the function defined by*

$$T(p(x)) = p'(x) + p(0)x^3.$$

1. *Show that* T *is a linear transformation.*
2. *Write down the matrix of* T *with respect to this basis and use it to find* $T(x^3 + x + 2)$.

Part 1. is left for you (do go through it!). The first step here is to figure out what T is "really" doing - p′(x) means the derivative of $p(x)$ and $p(0)$ is the rational number obtained by evaluating p at 0, which is just the constant term of $p(x)$. For part 2., we need to write the image under T of each of the basis elements, then write their coordinate vectors with respect to our basis.

1. $T(x^3 + 1) = 3x^2 + x^3 \to \begin{pmatrix} 1 \\ 3 \\ 0 \\ -4 \end{pmatrix}.$
2. $T(x^2 + 1) = 2x + x^3 \to \begin{pmatrix} 1 \\ 0 \\ 2 \\ -3 \end{pmatrix}.$
3. $T(x + 1) = 1 + x^3 \to \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$
4. $T(1) = x^3 \to \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix}.$

So the matrix of T with respect to this basis is

$$A = \left(\begin{array}{r|r|r|r} 1 & 1 & 1 & 1 \\ 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ -4 & -3 & 0 & -1 \end{array} \right).$$

(Note the lines between columns here are only to emphasize the meaning of the respective columns, they wouldn't normally be written). The coordinate vector of x^3+x+2 is $\begin{pmatrix}1\\0\\1\\0\end{pmatrix}$ and the coordinate vector of $T(x^3+x+2)$ is given by the matrix-vector product

$$\left(\begin{array}{r|r|r|r} 1 & 1 & 1 & 1 \\ 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ -4 & -3 & 0 & -1 \end{array}\right)\begin{pmatrix}1\\0\\1\\0\end{pmatrix} = 1\begin{pmatrix}1\\3\\0\\-4\end{pmatrix} + 0\begin{pmatrix}1\\0\\2\\-3\end{pmatrix} + 1\begin{pmatrix}1\\0\\0\\0\end{pmatrix} + 0\begin{pmatrix}1\\0\\0\\-1\end{pmatrix} = \begin{pmatrix}2\\3\\0\\-4\end{pmatrix}.$$

Thus $T(x^3+x+2) = 2(x^3+1)+3(x^2+1)-4(1) = 2x^3+3x^2+1$.

Of course it can be verified directly from the definition of T that this is the correct answer. The point of this exercise is not to suggest that this choice of basis and corresponding matrix is the most convenient mechanism for evaluating T but just to demonstrate how the process works and what the role of the basis is. Maybe it already suggests that one choice of basis might be better than another for describing a particular linear transformation.

5 Some bases are better than others