Lecture 6: The matrix of a Linear Transformation Relative to a Basis

Special Notation for the Next Three Lectures

In the next three lectures we will use the caligraphic font for bases e.g. $\mathscr{B}$ and Roman for linear transformations and matrices e.g. B.

Let V be a vector space and $\mathscr{B} = \{b_1, \ldots, b_n\}$ be a basis for V. Then any v has unique coordinates $[v]_{\mathscr{B}} = (x_1, \ldots, x_n)$ relative to $\mathscr{B}$ defined by

$$v = \sum_{i=1}^{n} x_i b_i$$

In the computations that follow we will usually write $[v]_{\mathscr{B}}$ as a column vector

$$[v]_{\mathscr{B}} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

So we will use column vectors for computations.

The Matrix of a Linear Transformation

Now let V and W be vector spaces and $T : V \longrightarrow W$ be a linear transformation. Suppose $\mathscr{B} = \{b_1, \ldots, b_n\}$ is a basis for V and $\mathscr{C} = \{c_1, \ldots, c_m\}$ is a basis for W.

Definition

The matrix of the linear transformation T relative to the basis $\mathscr{B}$ and $\mathscr{C}$ and written $M(T)$ or ${}_{\mathscr{C}}[T]_{\mathscr{B}}$ is the $m \times n$ matrix (a_{ij}) given by

$$T(b_j) = \sum_{i=1}^{m} a_{ij} c_i, \quad 1 \leq j \leq n, \qquad (*)$$

We will use the complicated notation ${}_{\mathscr{C}}[T]_{\mathscr{B}}$ when need to apply the very tricky "change of basis formula for the matrix of a linear transformation" - see Lecture 9. Think of $\mathcal{B}$ as the "input basis" and $\mathcal{C}$ as the "output basis". When we don't have to deal with the problem of changing bases we will use the simpler notation $M(T)$.

The Physical Meaning of $(*)$

It is important to understand the physical meaning of $(*)$. The first column of A is the coordinates of $T(b_1)$ relative to $c_1, \ldots, c_m$; the second column of A is the coordinates of $T(b_2)$ relative to $c_1, \ldots, c_m$, etc.
So

$$A = \begin{array}{c} \quad T(b_1) \quad T(b_2) \quad \ldots \quad T(b_n) \\ \left(\begin{array}{cccc} \downarrow & \downarrow & \ldots & \downarrow \end{array} \right) \end{array}$$

or

$$A = \begin{array}{c} \quad [T(b_1)]_{\mathscr{C}} \quad [T(b_2)]_{\mathscr{C}} \quad \ldots \quad [T(b_n)]_{\mathscr{C}} \\ \left(\begin{array}{cccc} \downarrow & \downarrow & \ldots & \downarrow \end{array} \right) \end{array}$$

We will usually have $V = W$ and $\mathscr{B} = \mathscr{C}$. Then ${}_{\mathscr{B}}[T]_{\mathscr{B}}$ is a square matrix.
If there is only one basis present we will write $M(T)$ instead of ${}_{\mathscr{B}}[T]_{\mathscr{B}}$.

Problem

Let $\mathrm{Pol}_3(\mathbb{R})$ be the set of polynomial functions of degree less than or equal to 3. Let $\dfrac{\mathrm{d}}{\mathrm{d}x} : \mathrm{Pol}_3(\mathbb{R}) \longrightarrow \mathrm{Pol}_3(\mathbb{R})$ be differentiation. Compute the matrix ${}_{\mathscr{B}}\left[\dfrac{\mathrm{d}}{\mathrm{d}x}\right]_{\mathscr{B}} = M\left(\dfrac{\mathrm{d}}{\mathrm{d}x}\right)$ relative to $\mathscr{B} = \{x_1, x, x^2, x^3\}$.

Solution

$\dfrac{\mathrm{d}}{\mathrm{d}x}(1) =$ the zero polynomial $= (0, 0, 0, 0)$ so the first column of $M\left(\dfrac{\mathrm{d}}{\mathrm{d}x}\right)$ is

$$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

$\frac{\mathrm{d}}{\mathrm{d}x}(x) = 1 = (1, 0, 0, 0)$ so the second column of $M\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)$ is $\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$.

$\frac{\mathrm{d}}{\mathrm{d}x}(x^2) = 2x = (0)1 + 2(x) + 0(x^2) + 0(x^3) = (0, 2, 0, 0)$ so the third column of $M\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)$ is $\begin{pmatrix} 0 \\ 2 \\ 0 \\ 0 \end{pmatrix}$.

Finally $\frac{\mathrm{d}}{\mathrm{d}x}(x^3) = 3x^2 = (0, 0, 3, 0)$.

We obtain:

$$_{\mathscr{B}}\left[\frac{\mathrm{d}}{\mathrm{d}x}\right]_{\mathscr{B}} = M\left(\frac{\mathrm{d}}{\mathrm{d}x}\right) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

We now prove :

Theorem

Fix a basis $\mathscr{B} = (b_1, \ldots, b_n)$ *for* V *and a basis* $\mathscr{C} = (c_1, \ldots, c_m)$ *for W. Then the map from* $L(V, W)$ *to the space of m by n matrices* $M_{m \times n}(\mathbb{R})$ *which assigns the matrix* ${}_{\mathscr{C}}[T]_{\mathscr{B}} \in M_{m \times n}(\mathbb{R})$ *to* $T \in \mathrm{Hom}(V, W)$ *is 1:1 and onto.*

Before starting the proof we recall the Theorem from Lecture 5. We restate it as a lemma.

Lemma

Let $\mathscr{B} = (b_1, \ldots, b_n)$ *be as above. Let* $w_1, w_2, \cdots, w_n$ *be arbitrary vectors in* W*. Then there exists a unique linear transformation* $T \in L(V, W)$ *such that*

$$T(b_1) = w_1, T(b_2) = w_2, \cdots, T(b_n) = w_n.$$

Now we can prove the Theorem.

Proof.
1:1: Let $T_1, T_2 \in \text{Hom}(V, W)$. Suppose ${}_{\mathscr{C}}[T_1]_{\mathscr{B}} = {}_{\mathscr{C}}[T_2]_{\mathscr{B}} = (a_{ij})$. Then for $1 \leq j \leq n$,

$$T_1(b_j) = \sum_{i=1}^{n} a_{ij} c_i \text{ and } T_2(b_j) = \sum_{i=1}^{n} a_{ij} c_i$$

.
Thus, $T_1(b_j) = T_2(b_j)$ hence $T_1 = T_2$.

Onto: Let $(a_{ij}) \in M_{m \times n}(\mathbb{R})$. Then define $w_j = \sum_{i=1}^{n} a_{ij} c_i$, $1 \leq j \leq n$.
In Lecture 5 we proved that there exits T with $T(b_i) = w_i$, $1 \leq i \leq n$. □

Notation: Let $\bullet$ denote matrix multiplication. Then we have the important

Proposition (1)

Let U, V, W be vector spaces with basis $\mathscr{A} = \{a_1, \ldots, a_m\}$, $\mathscr{B} = \{b_1, \ldots, b_n\}$ and $\mathscr{C} = \{c_1, \ldots, c_p\}$ respectively. Let $T : U \longrightarrow V$ and $S : V \longrightarrow W$ be linear transformations. Then

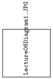
$$ {}_{\mathscr{C}}[S \circ T]_{\mathscr{A}} = {}_{\mathscr{C}}[S]_{\mathscr{B}} \bullet {}_{\mathscr{B}}[T]_{\mathscr{A}} $$

or much less precisely (see immediately below)

$$ M(S \circ T) = M(S) \circ M(T). $$

Proof. Put

$$ \begin{aligned} Z &= (z_{ik}) = M(S \circ T) = {}_{\mathscr{C}}[S \circ T]_{\mathscr{A}} \\ X &= (x_{ij}) = M(S) = {}_{\mathscr{C}}[S]_{\mathscr{B}} \\ Y &= (y_{jk}) = M(T) = {}_{\mathscr{B}}[T]_{\mathscr{A}} \end{aligned} $$



We will compute $(S \circ T)(a_k)$ in two ways.

The matrix (z_{ik}) is defined by

$$(S \circ T)(a_k) = \sum_{i=1}^{p} z_{ik} c_i.$$

Now we compute $(S \circ T)(a_k)$ another way. We have

$$(S \circ T)(a_k) = S(T(a_k)) \qquad (*)$$

But the matrix $Y = (y_{jk})$ is defined by

$$T(a_k) = \sum_{j=1}^{n} y_{jk} b_j. \qquad (**)$$

We substitute the RHS of $(**)$ into $(*)$ to get

$$\begin{aligned}(S \circ T)(a_k) &= S(T(a_k)) = S\left(\sum_{j=1}^{n} y_{jk}(b_j)\right) \\ &= \sum_{j=1}^{n} y_{jk} S\left(b_j\right). \qquad (\#)\end{aligned}$$

But the matrix $X = (x_{ij})$ is defined by

$$S(b_j) = \sum_{i=1}^{p} x_{ij} c_i. \qquad (\#\#)$$

We substitute $(\#\#)$ into $(\#)$ to get

$$\begin{aligned}
(S \circ T)(a_k) &= \sum_{j=1}^{n} y_{jk} \left(\sum_{i=1}^{p} x_{ij} c_i \right) \\
&= \sum_{j=1}^{n} \sum_{i=1}^{p} y_{jk} x_{ij} c_i \\
&= \sum_{i=1}^{p} \left(\sum_{j=1}^{n} x_{ij} y_{jk} \right) c_i.
\end{aligned}$$

Hence

$$\sum z_{ik} c_i = \sum_{i=1}^{p} \left(\sum_{j=1}^{n} x_{ij} y_{jk} \right) c_i .$$

Since c_i is a basis for W, we have

$$z_{ik} = \sum_{j=1}^{n} x_{ij} y_{jk} .$$

But the RHS is the ik-th entry of the product matrix $X \bullet Y$.

Remark: This wouldn't have worked if we had written the vectors $T(b_j)$ along the rows instead along the columns.

Proposition (2)

Let V be a vector space and $T \in L(V\,V) = \mathrm{Hom}\,(V,\,V)$. Let $\mathscr{B} = \{b_1,\,\ldots,\,b_n\}$ be a basis for V. Let $v \in V$. Then

$$[T(v)]_{\mathscr{B}} = {}_{\mathscr{B}}[T]_{\mathscr{B}}\,[v]_{\mathscr{B}}\,.$$

Proof. Suppose ${}_{\mathscr{B}}[T]_{\mathscr{B}} = A = (a_{ij})$ and $[v]_{\mathscr{B}} = (x_1,\,x_2,\,\ldots,\,x_n)$ so

$$v = \sum_{j=1}^{n} x_j b_j.$$

Then

$$T(v) = \sum_{j=1}^{n} x_j T(b_j). \qquad (*)$$

But by definition of the matrix ${}_{\mathscr{B}}[T]_{\mathscr{B}}$

$$T(b_j) = \sum_{i=1}^{n} a_{ij} b_i \qquad (**)$$

Subsitute $(**)$ into $(*)$ to obtain

$$\begin{aligned} T(v) &= \sum_{j=1}^{n} x_j \left(\sum_{i=1}^{n} a_{ij} b_i \right) \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} x_j a_{ij} b_i \\ &= \sum_{i=1}^{n} \left(\sum_{j=1}^{n} a_{ij} x_j \right) b_i \end{aligned}$$

Hence

$$[T(v)]_{\mathscr{B}} = \left(\sum_{j=1}^{n} a_{1j}x_j, \ldots, \sum_{j=1}^{n} a_{nj}x_j \right).$$

But

$$\begin{aligned} A\,[v]_{\mathscr{B}} &= \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ a_{21} & \ldots & a_{2n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \\ &= \begin{pmatrix} \sum_{j=1}^{n} a_{1j}x_j \\ \vdots \\ \sum_{j=1}^{n} a_{nj}x_j \end{pmatrix} \end{aligned}$$

At this stage we are not differentiating between row vectors and column vectors.

The Map from T to M(T)

We summarize what we have proved in this lecture.

Theorem (Text, Theorem 13.3)

Suppose $\dim V = n$ *and* $\{b_1, \ldots, b_n\}$ *is a basis for* V. *Then the map*

$$M : \mathrm{Hom}\,(V, V) \longrightarrow M_n(\mathbb{R})$$

that sends T *to* $M(T)$ *is 1:1, onto, linear and sends composition* $\circ$ *of linear transformations to multiplication* $\bullet$ *of matrices.*