7. LINEAR MAPS AND MATRICES

7.1 Representing linear maps with matrices

We saw examples of linear maps arising from multiplying by a matrix: for $A \in \mathcal{M}_{m\times n}(\mathbb{R})$, we defined $L_A \colon \mathbb{R}^n_{\text{col}} \to \mathbb{R}^m_{\text{col}}$ by $L_A(\mathbf{v}) = A\mathbf{v}$, and we defined $R_A \colon \mathbb{R}^m \to \mathbb{R}^n$ by $R_A(\mathbf{v}) = \mathbf{v}A$.

We shall see that linear maps are to matrices, as vectors are to coordinate vectors. Importantly recall that a vector has different coordinates in different coordinate systems and that each choice of coordinates (or basis) associates coordinates with a vector. Similarly given a linear map $T \colon V \to W$ for each choice of coordinates (or bases) for V and W we will see that T is represented by a matrix; change your choice of bases and that matrix will change too!

Definition 169 *Let V be an n-dimensional vector space over $\mathbb{F}$ with an ordered basis $\mathcal{V}$ of vectors $v_1, \ldots, v_n$. Let W be an m-dimensional vector space over $\mathbb{F}$ with an ordered basis $\mathcal{W}$ of vectors $w_1, \ldots, w_m$. So every vector in V and W is represented by a coordinate vector in $\mathbb{F}^n$ and $\mathbb{F}^m$ respectively.*

Let $T \colon V \to W$ be a linear transformation. We define the ***matrix for T with respect to the bases $\mathcal{V}$ and $\mathcal{W}$*** *to be the matrix which takes the coordinate vector of v to the coordinate vector of Tv.*

More explicitly, this is the $m \times n$ matrix $A = (a_{ij})$ where

$$T(v_i) = \sum_{k=1}^{m} a_{ki} w_k.$$

We will write ${}_{\mathcal{W}}T_{\mathcal{V}}$ for this matrix A.

Remark 170 *Firstly note that this matrix is well-defined. For each $1 \leqslant i \leqslant n$ then $T(v_i)$ can be uniquely expressed as a linear combination of $\mathcal{W}$.*

Remark 171 *Further $a_{1i}, \ldots, a_{mi}$ are the coordinates of $T(v_i)$. These are the entries in the ith column of A. The coordinate column vector of v_i is $\mathbf{e}_i^T$ and the ith column of A is $A\mathbf{e}_i^T$. So we can see that the entries of A are the coordinates of the images of the basis $\mathcal{V}$ as claimed.*

Note that this is what matrices normally do! Given an $m \times n$ matrix A then the first column of A equals $A\mathbf{e}_1$ where $\mathbf{e}_1 = (1, 0, \ldots, 0)$, and more generally the ith column is $A\mathbf{e}_i^T$.

Remark 172 *So if we use $\mathcal{V}$ and $\mathcal{W}$ to identify V and W with $\mathbb{F}^n_{\text{col}}$ and $\mathbb{F}^m_{\text{col}}$ then we identify T with L_A.*

Remark 173 *Importantly in this the bases are listed in an order. If the order of either basis changed then the matrix will change too.*

Remark 174 *If $V = W$ and we use the same ordered basis for both domain and codomain of $T\colon V \to V$, then we talk about the* ***matrix for T with respect to this basis****.*

Example 175 *Let $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ be defined by $T(x,y,z) = (0,x,y)$. This is linear (check!). If we take*

$$\mathbf{i} = (1,0,0), \qquad \mathbf{j} = (0,1,0), \qquad \mathbf{k} = (0,0,1),$$

as the basis $\mathcal{E}$ for both the domain and codomain then we see that

$$T(1,0,0) = (0,1,0), \qquad T(0,1,0) = (0,0,1), \qquad T(0,0,1) = (0,0,0).$$

Hence

$$_{\mathcal{E}}T_{\mathcal{E}} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Note that $({}_{\mathcal{E}}T_{\mathcal{E}})^2 = {}_{\mathcal{E}}T^2_{\mathcal{E}}$ and $({}_{\mathcal{E}}T_{\mathcal{E}})^3 = 0 = {}_{\mathcal{E}}T^3_{\mathcal{E}}$ (again check!).

Example 176 *Let $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ be defined by $T(x,y,z) = (0,x,y)$ as in the previous example. Now let $\mathcal{F}$ be the ordered basis*

$$\mathbf{f}_1 = (1,1,1), \qquad \mathbf{f}_2 = (1,1,0), \qquad \mathbf{f}_3 = (0,1,1).$$

Find (i) ${}_{\mathcal{E}}T_{\mathcal{F}}$, (ii) ${}_{\mathcal{F}}T_{\mathcal{E}}$ and (iii) ${}_{\mathcal{F}}T_{\mathcal{F}}$.

Solution. (i) Note that

$$T(1,1,1) = (0,1,1), \qquad T(1,1,0) = (0,1,1), \qquad T(0,1,1) = (0,0,1).$$

Hence

$$_{\mathcal{E}}T_{\mathcal{F}} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

(ii) Note that

$$T(1,0,0) = (0,1,0) = -\mathbf{f}_1 + \mathbf{f}_2 + \mathbf{f}_3, \qquad T(0,1,0) = (0,0,1) = \mathbf{f}_1 - \mathbf{f}_2, \qquad T(0,0,1) = (0,0,0).$$

Hence

$$_{\mathcal{F}}T_{\mathcal{E}} = \begin{pmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

(iii) Note that

$$T(1,1,1) = (0,1,1) = \mathbf{f}_3, \qquad T(1,1,0) = (0,1,1) = \mathbf{f}_3, \qquad T(0,1,1) = (0,0,1) = \mathbf{f}_1 - \mathbf{f}_2.$$

Hence

$$_{\mathcal{F}}T_{\mathcal{F}} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & -1 \\ 1 & 1 & 0 \end{pmatrix}.$$

■

Proposition 177 *Let V be an n-dimensional vector space over $\mathbb{F}$ with ordered basis $\mathcal{V}$. Let W be an m-dimensional vector space over $\mathbb{F}$ with ordered basis $\mathcal{W}$. Then*
(i) the matrix ${}_{\mathcal{W}}0_{\mathcal{V}}$ of the zero map is $0_{m\times n}$ for any choice of $\mathcal{V}$ and $\mathcal{W}$;
(ii) the matrix ${}_{\mathcal{V}}id_{\mathcal{V}}$ of identity map I_n for any choice of $\mathcal{V}$;
(iii) if $S\colon V \to W$, $T\colon V \to W$ are linear and α, $\beta \in \mathbb{F}$, then

$$ {}_{\mathcal{W}}(\alpha S + \beta T)_{\mathcal{V}} = \alpha\left({}_{\mathcal{W}}S_{\mathcal{V}}\right) + \beta\left({}_{\mathcal{W}}T_{\mathcal{V}}\right). $$

Proof. These are left as exercises. ■

Theorem 178 *Let U, V, W be finite-dimensional vector spaces over $\mathbb{F}$, of dimensions m, n, p, with ordered bases $\mathcal{U}$, $\mathcal{V}$, $\mathcal{W}$ respectively. Let $S\colon U \to V$ and $T\colon V \to W$ be linear. Let*

$$ A = {}_{\mathcal{V}}S_{\mathcal{U}} \quad \text{and} \quad B = {}_{\mathcal{W}}T_{\mathcal{V}}. $$

Then

$$ BA = {}_{\mathcal{W}}TS_{\mathcal{U}}. $$

Proof. Note that A is $n \times m$ and B is $p \times n$, so the product matrix BA is $p \times m$. Let $\mathcal{U}$ be $u_1, \ldots, u_m$, $\mathcal{V}$ be $v_1, \ldots, v_n$ and $\mathcal{W}$ be $w_1, \ldots, w_p$.

As usual, we write $A = (a_{ij})$ and $B = (b_{ij})$. By definition of A and B, we have

$$ \begin{aligned} S(u_i) &= \sum_{j=1}^{n} a_{ji} v_j \text{ for } 1 \leqslant i \leqslant m; \\ T(v_j) &= \sum_{k=1}^{p} b_{kj} w_k \text{ for } 1 \leqslant j \leqslant n. \end{aligned} $$

Now for $1 \leqslant i \leqslant m$ we have

$$ \begin{aligned} (T \circ S)(u_i) = T(S(u_i)) &= T\left(\sum_{j=1}^{n} a_{ji} v_j\right) \\ &= \sum_{j=1}^{n} a_{ji} T(v_j) \text{ since } T \text{ is linear} \\ &= \sum_{j=1}^{n} a_{ji} \sum_{k=1}^{p} b_{kj} w_k \\ &= \sum_{k=1}^{p} \left(\sum_{j=1}^{n} b_{kj} a_{ji}\right) w_k \\ &= \sum_{k=1}^{p} (BA)_{ki} w_k. \end{aligned} $$

Thus, by definition, ${}_{\mathcal{W}}TS_{\mathcal{U}} = BA$. ■

Remark 179 *This is why we define multiplication of matrices in the way that we do!*

Remark 180 *As we are about to see, this gives a relatively clear and painless proof that matrix multiplication is associative as composition is associative.*

Corollary 181 *Take $A \in \mathcal{M}_{m\times n}(\mathbb{F})$, take $B \in \mathcal{M}_{n\times p}(\mathbb{F})$, take $C \in \mathcal{M}_{p\times q}(\mathbb{F})$. Then $A(BC) = (AB)C$.*

Proof. We consider the left multiplication maps

$$L_A \colon \mathbb{F}^n_{\text{col}} \to \mathbb{F}^m_{\text{col}} \quad \text{and} \quad L_B \colon \mathbb{F}^p_{\text{col}} \to \mathbb{F}^n_{\text{col}} \quad \text{and} \quad L_C \colon \mathbb{F}^q_{\text{col}} \to \mathbb{F}^p_{\text{col}}.$$

With respect to the standard bases of these spaces, the matrices of L_A, L_B, L_C are A, B, C respectively.Hence, by the previous theorem $A(BC)$ and $(AB)C$ are the matrices of

$$L_A \circ (L_B \circ L_C) \colon \mathbb{F}^q_{\text{col}} \to \mathbb{F}^m_{\text{col}}, \qquad \text{and} \quad (L_A \circ L_B) \circ L_C \colon \mathbb{F}^q_{\text{col}} \to \mathbb{F}^m_{\text{col}}$$

respectively. But composition of functions is associative, so

$$L_A \circ (L_B \circ L_C) = (L_A \circ L_B) \circ L_C$$

and hence $A(BC) = (AB)C$. ■

Corollary 182 *Let V be a finite-dimensional vector space and let $T \colon V \to V$ be an invertible linear transformation. Let A be the matrix of T with respect to an ordered basis (for both domain and codomain). Then A is invertible, and A^{-1} is the matrix of T^{-1} with respect to the same basis.*

Proof. Exercise. ■

7.2 Change of basis

Question Take two matrices for the same linear transformation with respect to different bases. How are the matrices related?

Example 183 *Define $T \colon \mathbb{R}^2 \to \mathbb{R}^2$ by $T(x, y) = (2x + y, 3x - 2y)$. To find the matrix of T with respect to the standard ordered basis $\mathcal{E}$, note that*

$$T(1,0) = (2,3) \qquad \textit{and} \qquad T(0,1) = (1,-2)$$

so the matrix for T with respect to this basis is

$$_{\mathcal{E}}T_{\mathcal{E}} = \begin{pmatrix} 2 & 1 \\ 3 & -2 \end{pmatrix}.$$

That is $T = L_A$. Let $\mathbf{f}_1 = (1,-2)$ and $\mathbf{f}_2 = (-2,5)$. Then $\mathbf{f}_1$, $\mathbf{f}_2$ is an ordered basis of $\mathbb{R}^2$ which we will denote as $\mathcal{F}$. Note that

$$\begin{aligned} T(\mathbf{f}_1) &= (0,7) = 14\mathbf{f}_1 + 7\mathbf{f}_2 \\ T(\mathbf{f}_2) &= (1,-16) = -27\mathbf{f}_1 - 14\mathbf{f}_2 \end{aligned}$$

so the matrix for T with respect to this basis is

$$_{\mathcal{F}}T_{\mathcal{F}} = \begin{pmatrix} 14 & -27 \\ 7 & -14 \end{pmatrix}.$$

How are these two matrices related? Well, by Theorem 178, we can see that

$$_{\mathcal{F}}T_{\mathcal{F}} = (_{\mathcal{F}}I_{\mathcal{E}})\,(_{\mathcal{E}}T_{\mathcal{E}})\,(_{\mathcal{E}}I_{\mathcal{F}})\,.$$

The matrix $_{\mathcal{E}}I_{\mathcal{F}}$ represents the identity transformation, so it is does not change vectors; however it changes the coordinate vector for a vector with respect to some basis $\mathcal{F}$ to the coordinate vector for the ***same*** *vector with respect to a* ***different*** basis *$\mathcal{E}$. Note that the inverse of this matrix is $_{\mathcal{F}}I_{\mathcal{E}}$.*

We can take $\mathbf{f}_1$, $\mathbf{f}_2$ and write them with respect to $\mathbf{e}_1$, $\mathbf{e}_2$: we have

$$\mathbf{f}_1 = \mathbf{e}_1 - 2\mathbf{e}_2, \qquad \mathbf{f}_2 = -2\mathbf{e}_1 + 5\mathbf{e}_2$$

so we get a 'change of basis matrix'

$$_{\mathcal{E}}I_{\mathcal{F}} = \begin{pmatrix} 1 & -2 \\ -2 & 5 \end{pmatrix}.$$

If this matrix is applied to $(1,0)^T$ then this coordinate vector represents $\mathbf{f}_1$. The image of the coordinate vector $(1,0)^T$ is $(1,-2)^T$ which represents $\mathbf{e}_1 - 2\mathbf{e}_2$. But this is of course the same vector! This vector just has different coordinates with respect to the bases $\mathcal{E}$ and $\mathcal{F}$.

It is then the case that

$$_{\mathcal{F}}I_{\mathcal{E}} = \begin{pmatrix} 1 & -2 \\ -2 & 5 \end{pmatrix}^{-1} = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix},$$

which represents that

$$\mathbf{e}_1 = 5\mathbf{f}_1 + 2\mathbf{f}_2, \qquad \mathbf{e}_2 = 2\mathbf{f}_1 + \mathbf{f}_2.$$

And we can verify that

$$\begin{aligned} & (_{\mathcal{F}}I_{\mathcal{E}})\,(_{\mathcal{E}}T_{\mathcal{E}})\,(_{\mathcal{E}}I_{\mathcal{F}}) \\ = & \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 3 & -2 \end{pmatrix}\begin{pmatrix} 1 & -2 \\ -2 & 5 \end{pmatrix} \\ = & \begin{pmatrix} 16 & 1 \\ 7 & 0 \end{pmatrix}\begin{pmatrix} 1 & -2 \\ -2 & 5 \end{pmatrix} \\ = & \begin{pmatrix} 14 & -27 \\ 7 & -14 \end{pmatrix} = {}_{\mathcal{F}}T_{\mathcal{F}} \end{aligned}$$

as expected.

Corollary 184 ***(Change of basis theorem)*** *Let V be a finite-dimensional vector space over $\mathbb{F}$ with ordered bases $\mathcal{V}$, $\mathcal{V}'$. Let W be a finite-dimensional vector space over $\mathbb{F}$ with ordered bases $\mathcal{W}$, $\mathcal{W}'$. Let $T\colon V \to W$ be a linear map. Then*

$$_{\mathcal{W}'}T_{\mathcal{V}'} = (_{\mathcal{W}'}I_{\mathcal{W}})\,(_{\mathcal{W}}T_{\mathcal{V}})\,(_{\mathcal{V}}I_{\mathcal{V}'})\,.$$

Proof. This is an immediate corollary to Theorem 178. ■

Corollary 185 *(**Change of basis theorem version 2**) Let V be a finite-dimensional vector space over $\mathbb{F}$ with ordered bases $\mathcal{V}$, $\mathcal{V}'$ and let $T\colon V \to V$ be a linear map. Then*

$$_{\mathcal{V}'}T_{\mathcal{V}'} = ({}_{\mathcal{V}'}I_{\mathcal{V}})\,({}_{\mathcal{V}}T_{\mathcal{V}})\,({}_{\mathcal{V}}I_{\mathcal{V}'})\,.$$

If we set $A = {}_{\mathcal{V}'}T_{\mathcal{V}'}, B = {}_{\mathcal{V}}T_{\mathcal{V}}$ and $P = {}_{\mathcal{V}}I_{\mathcal{V}'}$ then note

$$A = P^{-1}BP.$$

Proof. This is a special case of the previous corollary. ■

Definition 186 *Take A, $B \in \mathcal{M}_{n\times n}(\mathbb{F})$. If there is an invertible $n \times n$ matrix P such that $A = P^{-1}BP$, then we say that A and B are **similar**. Similarity is then an equivalence relation.*

Remark 187 *So two matrices representing the same linear transformation from a finite-dimensional vector space to itself, but with respect to different bases, are similar.*

Remark 188 ***Properties of Linear Maps** As many different matrices can represent the **same** linear transformation $T\colon V \to V$ it would be disturbing if different conclusions about the properties of T could be determined by using different matrix representatives. For example, if we said a linear map T is invertible if a matrix representative of it is invertible, could T end up being invertible and not invertible? Reassuringly the answer is no.*

Let $A = {}_{\mathcal{V}}T_{\mathcal{V}}$ and $B = {}_{\mathcal{W}}T_{\mathcal{W}}$ be matrices representing T with respect to two bases, so that $A = P^{-1}BP$ for some invertible P. Then

- *A is invertible if and only if B is invertible.*
- *The trace of A equals the trace of B. [This follows from the identity* $\operatorname{trace}(MN) = \operatorname{trace}(NM)$.*]*
- *A functional identity satisfied by A, such as $A^2 = A$, is also satisfied by B.*
- *The determinant of A equals the determinant of B. [Determinants will be formally defined in Linear Algebra II next term.]*
- *The eigenvalues of A equal the eigenvalues of B. [Eigenvalues will be formally defined in Linear Algebra II next term.]*

Thus we may, in a well-defined fashion, refer to the invertibility, trace, determinant of a linear map.

Note we cannot, in a well-defined manner, refer to the transpose of a linear map. If $A = P^{-1}BP$ then it need not be the case that $A^T = P^{-1}B^TP$. Though you might note that this is true if P is orthogonal! (This is something that will be addressed when you meet adjoints in the second year.)

7.3 Matrices and rank

For a matrix $A \in \mathcal{M}_{m\times n}(\mathbb{F})$, we have defined the row space and row rank, and analogously the column space and column rank. It makes sense to ask if $\text{rowrank}(A)$ and $\text{colrank}(A)$ related?

Remark 189 *From the definitions, we see that* $\text{Col}(A) = \text{Row}(A^T)$ *and so* $\text{colrank}(A) = \text{rowrank}(A^T)$*. Similarly,* $\text{Row}(A) = \text{Col}(A^T)$ *and so* $\text{rowrank}(A) = \text{colrank}(A^T)$*.*

We first prove the following:

Lemma 190 *The linear system* $(A|\mathbf{b})$ *is consistent if and only if* $\text{Col}(A|\mathbf{b}) = \text{Col}(A)$.

Proof. Say that A is $m \times n$ and denote the columns of A as $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n$. Then

$$\begin{aligned} (A|\mathbf{b}) \text{ is consistent} &\iff A\mathbf{x} = \mathbf{b} \text{ for some } \mathbf{x} \in \mathbb{F}^n_{\text{col}} \\ &\iff x_1\mathbf{c}_1 + \cdots x_n\mathbf{c}_n = \mathbf{b} \text{ for some } \mathbf{x} \in \mathbb{F}^n_{\text{col}} \\ &\iff \mathbf{b} \in \text{Col}(A) \\ &\iff \text{Col}(A|\mathbf{b}) = \text{Col}(A). \end{aligned}$$

■

Theorem 191 *The column rank of a matrix equals its row rank.*

Proof. We prove this by induction on the number of columns in the matrix. A non-zero $m \times 1$ matrix has column rank 1 and also row rank 1 as the matrix reduces to $\mathbf{e}_1^T$; the column rank and row rank of $0_{m\times 1}$ are both 0. So the $n = 1$ case is true. Suppose, as our inductive hypothesis, that column rank and row rank are equal for $m \times n$ matrices. Any $m \times (n+1)$ matrix $(A|\mathbf{b})$ can be considered as an $m \times n$ matrix A alongside $\mathbf{b}$ in $\mathbb{F}^m_{\text{col}}$. If the system $(A|\mathbf{b})$ is consistent then

$$\begin{aligned} \text{colrank}(A|\mathbf{b}) &= \text{colrank}(A) && \text{[by previous lemma]} \\ &= \text{rowrank}(A) && \text{[by inductive hypothesis]} \\ &= \text{rowrank}(A|\mathbf{b}) && \text{[see Remark 45].} \end{aligned}$$

On the other hand, if the system $(A|\mathbf{b})$ has no solutions then

$$\begin{aligned} \text{colrank}(A|\mathbf{b}) &= (\text{colrank}A) + 1 && \text{as } \mathbf{b} \notin \text{Col}(A) \\ &= (\text{rowrank}A) + 1 && \text{[by inductive hypothesis]} \\ &= \text{rowrank}(A|\mathbf{b}) && \text{[see Remark 45].} \end{aligned}$$

So if the system is consistent the row rank and column rank maintain their common value. If inconsistent, then $\mathbf{b}$ adds a further dimension to the column space and $(0\,0 \cdots 0 \,|\, 1)$ adds an extra dimension to the row space. Either way the column rank and row rank of $(A|\mathbf{b})$ still agree and the proof follows by induction. ■

We provide here a second proof of the result, as it takes a somewhat different approach.

Theorem 192 *Let $P = QR$ where P, Q, R are respectively $k \times l$, $k \times m$, $m \times l$ matrices over the same field.*
(a) Then the row rank of P is at most m.
(b) Let $p = \mathrm{rowrank} P$. Then P may be written as the product of a $k \times p$ matrix and a $p \times l$ matrix.
(c) The row rank and column rank of a matrix are equal.

Solution. (a) By Proposition 117(c) we have

$$\mathrm{Row}(P) = \mathrm{Row}(QR) \leqslant \mathrm{Row}(R).$$

Hence

$$\mathrm{rowrank}(P) = \dim \mathrm{Row}(P) \leqslant \dim \mathrm{Row}(R) = \mathrm{rowrank}(R) \leqslant m.$$

(b) There is a $k \times k$ invertible matrix E such that $EP = \mathrm{RRE}(P)$ and hence

$$P = E^{-1}\mathrm{RRE}(P).$$

Now let $\tilde{E}$ denote the first p columns of E^{-1} and $\tilde{P}$ denote the first p rows of $\mathrm{RRE}(P)$. As the last $k - p$ rows of $\mathrm{RRE}(P)$ are zero rows, we still have

$$P = \tilde{E}\tilde{P},$$

where $\tilde{E}$ is a $k \times p$ and $\tilde{P}$ is a $p \times l$ matrix.

(c) From (a) and (b) we know that the row rank of a $k \times l$ matrix P is the minimal value p such that P can be written as the product QR of a $k \times p$ matrix and a $p \times l$ matrix. Whenever $P = QR$ then

$$\underset{l \times k}{P^T} = \underset{l \times p}{R^T}\,\underset{p \times k}{Q^T}.$$

So the row rank of P^T is similarly p. But the $\mathrm{rowrank} P^T = \mathrm{colrank} P$ as required. ■