LECTURE 10

Matrix representations (1)

10.1 Matrix representations

By the rank-nullity formula, if the kernel of a map $f : V \to W$ is trivial, then the dimension of the image is the same as that of the domain. Since linearly independent sets in V are mapped to linearly independent sets in $f(V)$, the action of f on a basis of V will yield a basis for $f(V)$. Thus, the action of f on a vector $\mathbf{v} \in V$ can be reduced to the action of f on basis elements. Let us consider an example to see how this works.

Example 10.1 Let $f : \mathcal{P}_2(\mathbb{R}) \to \mathcal{P}_1(\mathbb{R})$ be given by

$$f(a_0 + a_1 x + a_2 x^2) = a_2 + (a_0 + a_1 + a_2)x.$$

Consider the following bases for $V = \mathcal{P}_2(\mathbb{R})$ and $W = \mathcal{P}_1(\mathbb{R})$, respectively,

$$B_V = \{p_1, p_2, p_3\} := \left\{1, 1 + x, 1 + x + x^2\right\}, \qquad B_W = \{q_1, q_2\} := \{1, x\}.$$

The action of f on each basis element of $\mathcal{P}_2$ is as follows:

$$\begin{cases} f(p_1) = f(1) & = x = 0 \cdot q_1 + 1 \cdot q_2, \\ f(p_2) = f(1 + x) & = 2x = 0 \cdot q_1 + 2 \cdot q_2, \\ f(p_3) = f(1 + x + x^2) & = 3x + 1 = 1 \cdot q_1 + 3 \cdot q_2. \end{cases}$$

Thus, we can establish the correspondence

$$p_1 \longleftrightarrow \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \xrightarrow{f} \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad p_2 \longleftrightarrow \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \xrightarrow{f} \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \quad p_3 \longleftrightarrow \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \xrightarrow{f} \begin{bmatrix} 1 \\ 3 \end{bmatrix},$$

where we used the convention that the coordinates in the bases B_V, B_W are displayed in column vectors. We can therefore summarise the action of f on the basis of V as follows

$$[p_1, p_2, p_3] \longleftrightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{f} \begin{bmatrix} 0 & 0 & 1 \\ 1 & 2 & 3 \end{bmatrix}.$$

On the other hand, any $p \in \mathcal{P}_2(\mathbb{R})$ will be mapped by f as follows:

$$f(p) = f(c_1p_1 + c_2p_2 + c_3p_3) = c_1f(p_1) + c_2f(p_2) + c_3f(p_3)$$

$$= c_1x + c_2(2x) + c_3(3x + 1) = c_3q_1 + (c_1 + 2c_2 + 3c_3)q_2 = q.$$

Hence, the action of mapping p to q can be viewed as the action of mapping the coordinates of p to the coordinates of q via a rectangular matrix:

$$\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} \xrightarrow{\begin{bmatrix} 0 & 0 & 1 \\ 1 & 2 & 3 \end{bmatrix}} \begin{bmatrix} c_3 \\ c_1 + 2c_3 + 3c_3 \end{bmatrix},$$

or equivalently,

$$f(p) \longleftrightarrow \begin{bmatrix} 0 & 0 & 1 \\ 1 & 2 & 3 \end{bmatrix} * \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} \longleftrightarrow \begin{bmatrix} c_3 \\ c_1 + 2c_3 + 3c_3 \end{bmatrix} \longleftrightarrow q.$$

The operation $*$ can be immediately seen as the matrix-vector product between the matrix associated with f and the coordinates of a vector in the domain of f. In fact, this example represents *the motivation for the definition of the matrix-product as we know it.* It is also the motivation for the definition of a matrix representation of a linear map – included below.

For simplicity of presentation, in the following we consider only vector spaces over $\mathbb{R}$.

Definition 10.1 Let V, W be finite-dimensional vector spaces over $\mathbb{R}$. Let $f : V \to W$ be a linear map and assume that

$$B_V = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}, \qquad B_W = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$$

are basis sets for V and W, respectively. For $i = 1, 2, \ldots, n$, let

$$f(\mathbf{v}_i) = \sum_{j=1}^{m} a_{ji}\mathbf{w}_j.$$

The matrix $A := [a_{ji}] \in \mathbb{R}^{m\times n}$ is **the matrix representation of f relative to the bases** B_V, B_W.

The above definition is validated by the following general result.

Proposition 10.1 Let $f \in \mathcal{L}(V, W)$ and let $A \in \mathbb{R}^{m\times n}$ be its matrix representation relative to bases

$$B_V = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}, \qquad B_W = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}.$$

Let

$$\mathbf{v} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_n\mathbf{v}_n = \sum_{i=1}^{n} x_i\mathbf{v}_i \in V$$

and

$$f(\mathbf{v}) = y_1\mathbf{w}_1 + y_2\mathbf{w}_2 + \cdots + y_m\mathbf{w}_m = \sum_{j=1}^{m} y_j\mathbf{w}_j \in W.$$

Then $\mathbf{y} = A\mathbf{x}$, where $y_j = [\mathbf{y}]_j$ for $j = 1, 2, \ldots, m$ and $x_i = [\mathbf{x}]_i$ for $i = 1, 2, \ldots, n$.

Proof. We have

$$f(\mathbf{v}) = f\left(x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \ldots + x_n\mathbf{v}_n\right) = \sum_{j=1}^{n} x_j f(\mathbf{v}_j) = \sum_{j=1}^{n} x_j \sum_{i=1}^{m} a_{ij}\mathbf{w}_i = \sum_{i=1}^{m}\sum_{j=1}^{n} x_j a_{ij}\mathbf{w}_i = \sum_{i=1}^{m} y_i\mathbf{w}_i,$$

where, for $i = 1, 2, \ldots, m$,

$$y_i := \sum_{j=1}^{n} x_j a_{ij} = \sum_{j=1}^{n} a_{ij}x_j = [A\mathbf{x}]_i.$$

■

The above statement has the following counterpart.

Proposition 10.2 Let $A \in \mathbb{R}^{m\times n}$. Then A is the matrix representation of some linear map relative to any bases for any vector spaces of dimensions n and m, respectively.

Proof. Let V, W be arbitrary vector spaces of dimensions n and m, with arbitrary bases

$$B_V = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}, \qquad B_W = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}.$$

Then any $\mathbf{v} \in V$ can be represented as

$$\mathbf{v} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_n\mathbf{v}_n = \sum_{j=1}^{n} x_j\mathbf{v}_j = \sum_{j=1}^{n} [\mathbf{x}]_j\mathbf{v}_j, \qquad [\mathbf{x}]_j := x_j \ (j = 1, \ldots, n).$$

Define $\mathbf{y} := A\mathbf{x}$, i.e.,

$$y_i := [\mathbf{y}]_i := [A\mathbf{x}]_i \ (i = 1, \ldots, m).$$

Since $\mathbf{v}$ is arbitrary in V, we can define a map $f : V \to W$ via

$$f(\mathbf{v}) := y_1\mathbf{w}_1 + y_2\mathbf{w}_2 + \cdots + y_m\mathbf{w}_m = \sum_{i=1}^{m} [\mathbf{y}]_i\mathbf{w}_i.$$

Claim: f is a linear map. To see this, let $\mathbf{v}$ be as above and choose another vector $\tilde{\mathbf{v}} \in V$ represented as

$$\tilde{\mathbf{v}} = \tilde{x}_1\mathbf{v}_1 + \tilde{x}_2\mathbf{v}_2 + \ldots + \tilde{x}_n\mathbf{v}_n = \sum_{j=1}^{n} \tilde{x}_j\mathbf{v}_j = \sum_{j=1}^{n} [\tilde{\mathbf{x}}]_j\mathbf{v}_j, \qquad [\tilde{\mathbf{x}}]_j := \tilde{x}_j \ (j = 1, \ldots, n).$$

Define $\tilde{\mathbf{y}} = A\tilde{\mathbf{x}}$, i.e.,

$$\tilde{y}_i := [\tilde{\mathbf{y}}]_i := [A\tilde{\mathbf{x}}]_i \ (i = 1, \ldots, m).$$

Then, by the above definition of f,

$$f(\tilde{\mathbf{v}}) = \tilde{y}_1\mathbf{w}_1 + \tilde{y}_2\mathbf{w}_2 + \cdots + \tilde{y}_m\mathbf{w}_m = \sum_{i=1}^{m} [\tilde{\mathbf{y}}]_i\mathbf{w}_i.$$

We have

$$af(\mathbf{v}) + bf(\tilde{\mathbf{v}}) = a\left(\sum_{i=1}^{m} [\mathbf{y}]_i\mathbf{w}_i\right) + b\left(\sum_{i=1}^{m} [\tilde{\mathbf{y}}]_i\mathbf{w}_i\right) = \sum_{i=1}^{m} [a\mathbf{y} + b\tilde{\mathbf{y}}]_i\mathbf{w}_i.$$

On the other hand,

$$a\mathbf{v} + b\tilde{\mathbf{v}} = a\left(\sum_{j=1}^{n} [\mathbf{x}]_j\mathbf{v}_j\right) + b\left(\sum_{j=1}^{n} [\tilde{\mathbf{x}}]_j\mathbf{v}_j\right) = \sum_{j=1}^{n} [a\mathbf{x} + b\tilde{\mathbf{x}}]_j\mathbf{v}_j.$$

Since $A(a\mathbf{x} + b\tilde{\mathbf{x}}) = a\mathbf{y} + b\tilde{\mathbf{y}}$, we find

$$f(a\mathbf{v} + b\tilde{\mathbf{v}}) = \sum_{i=1}^{m} [a\mathbf{y} + b\tilde{\mathbf{y}}]_i\mathbf{w}_i = af(\mathbf{v}) + bf(\tilde{\mathbf{v}}).$$

Therefore, f is a linear map. By definition, A is the matrix representation of f relative to B_V, B_W. ■

We note the following uniqueness result.

Proposition 10.3 Let $f \in \mathcal{L}(V, W)$. Then its matrix representation relative to bases B_V, B_W is unique.

Proof. Assume that there are two distinct matrix representations A, A' relative to bases

$$B_V = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}, \qquad B_W = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}.$$

By definition, for all $j = 1, 2, \ldots, n$,

$$f(\mathbf{v}_j) = \sum_{i=1}^{m} a_{ij}\mathbf{w}_i = \sum_{i=1}^{m} a'_{ij}\mathbf{w}_i.$$

Taking the difference, we find

$$\mathbf{0}_W = \sum_{i=1}^{m} a_{ij}\mathbf{w}_i - \sum_{i=1}^{m} a'_{ij}\mathbf{w}_i = \sum_{i=1}^{m} \tilde{a}_{ij}\mathbf{w}_i,$$

where $\tilde{a}_{ij} = a_{ij} - a'_{ij}$. Therefore, for all $j = 1, 2, \ldots, n$, the zero vector $\mathbf{0}_W$ is a linear combination of basis elements; by linear independence, we have $\tilde{a}_{ij} = 0$ for all i, or $a_{ij} = a'_{ij}$ for all i, j, i.e., $A = A'$, which is a contradiction. ∎

We have now seen that matrices can be used to define linear maps and most certainly will always represent linear maps. When dealing with matrices in this sense, the following terminology is used.

Definition 10.2 — Column space of a matrix. Let $A \in \mathbb{R}^{n\times k}$. The column space of A, denoted by col A is the subspace of $\mathbb{R}^n$ spanned by the columns of A:

$$\text{col } A = \text{span}\{\mathbf{c}_1(A), \mathbf{c}_2(A), \ldots, \mathbf{c}_k(A)\}.$$

Thus, the column space is the image or range of A, when viewed as a linear map.

Definition 10.3 — Rank of a matrix. Let $A \in \mathbb{R}^{n\times k}$. The rank of A is the number of linearly independent columns of A.

Using our results on dimension, we immediately derive this result.

Proposition 10.4 Let $A \in \mathbb{R}^{n\times k}$. Then rank $A = \dim \text{col } A$.

It is evident that for any non-zero matrix in $\mathbb{R}^{n\times m}$ we have $1 \leq \text{rank } A \leq m$. This allows for the following terminology.

Definition 10.4 We say $A \in \mathbb{R}^{n\times m}$ has **full rank** if rank $A = m$. Otherwise, A is said to be **rank-deficient**.

Before we consider some examples, we include the following counterpart to Proposition 8.1 which stated that $\mathcal{L}(V, W)$ is a vector space.

Proposition 10.5 The set of matrices $\mathbb{R}^{m\times n}$ is a vector space over $\mathbb{R}$.

Proof. Exercise. ∎

This result confirms that the set of matrix representations for linear maps inherit the properties and structures afforded by the set of linear maps.

10.2 Examples

Let us consider some examples of derivations of matrices associated with linear maps. In the following, we assume that the domain and codomain are equipped with basis sets

$$B_V = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}, \qquad B_W = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m\}.$$

Example 10.2 — Zero map. Let $o : V \to W$ be the zero map with $A \in \mathbb{R}^{m\times n}$ as its matrix representation. Then, for all $j = 1, \dots, n$,

$$o(\mathbf{v}_j) = \sum_{i=1}^{m} a_{ij}\mathbf{w}_i = \mathbf{0}_W.$$

In other words, given any j, the zero vector $\mathbf{0}_W$ is a linear combination of basis elements; by linear independence, we have $a_{ij} = 0$ for all i. Hence A is the $m \times n$ zero matrix $A = O_{m,n}$.

Example 10.3 — Identity map. Let $id : V \to V$ be the identity map with $A \in \mathbb{R}^{m\times n}$ as its matrix representation relative to bases B_V and $B_W = B_V$. Then, for all $j = 1, \dots, n$,

$$id(\mathbf{v}_j) = \sum_{i=1}^{m} a_{ij}\mathbf{v}_i = \mathbf{v}_j \iff \sum_{i=1}^{m} \tilde{a}_{ij}\mathbf{v}_i = \mathbf{0}_V, \quad \text{where } \tilde{a}_{ij} = \begin{cases} a_{ij} & i \neq j, \\ a_{ij} - 1 & i = j. \end{cases}$$

Hence, given any j, the zero vector $\mathbf{0}_V$ is a linear combination of basis elements; by linear independence, we have $\tilde{a}_{ij} = 0$ for all i, so that for all j

$$a_{ij} = \begin{cases} 0 & i \neq j, \\ 1 & i = j. \end{cases}$$

Therefore, $A = I_n \in \mathbb{R}^{n\times n}$ is the identity matrix.

The next example considers a restriction of the differentiation map to spaces of polynomials. This ensures that the domain and codomain are finite-dimensional vector spaces.

Example 10.4 — Differentiation map on $\mathcal{P}_n(\mathbb{R})$. Let $D : \mathcal{P}_n(\mathbb{R}) \to \mathcal{P}_{n-1}(\mathbb{R})$ be given by $(Dp)(x) = p'(x)$. Note that $V = P_n(\mathbb{R})$, $W = P_{n-1}(R)$, with $\dim V = n + 1, \dim W = n$. Let

$$B_V = \left\{q_j = x^{j-1}, j = 1, 2, \dots, n+1\right\}, \qquad B_W = \left\{q_i = x^{i-1}, i = 1, 2, \dots, n\right\}.$$

Then the matrix representation $A \in \mathbb{R}^{n\times(n+1)}$ satisfies for all $j = 1, \dots, n+1$

$$D(q_j) = \sum_{i=1}^{n} a_{ij}q_i.$$

Let us consider the case $j = 1$ first:

$$j = 1 : 0 = D(q_1) = \sum_{i=1}^{n} a_{i1}q_i \implies a_{i1} = 0, \quad i = 1, 2, \dots, n.$$

For any $j > 1$, matching the monomials on the left and right, we find

$$D(q^j) = (j-1)x^{j-2} = \sum_{i=1}^{n} a_{ij}x^{i-1} \iff a_{ij} = \begin{cases} 0 & j-2 \neq i-1 \\ j-1 & j-2 = i-1 \end{cases} \iff a_{ij} = \begin{cases} 0 & i \neq j-1, \\ j-1 & i = j-1. \end{cases}.$$

The resulting matrix is included below.

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & 2 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 3 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 & n \end{bmatrix} \in \mathbb{R}^{n\times(n+1)}.$$

Our next example considers the anti-derivative/integration map. As for the previous example, to ensure that the domain and codomain are finite-dimensional, we restrict them to polynomial spaces.

Example 10.5 — Integration map on $\mathcal{P}_n(\mathbb{R})$. Let $V = \mathcal{P}_n(\mathbb{R})$. Define $W = \mathcal{P}^0_{n+1}(\mathbb{R}) = \{p \in \mathcal{P}_{n+1} : p(0) = 0\}$ and note that $W \leq \mathcal{P}_{n+1}(\mathbb{R})$. Let

$$B_V = \left\{q_j = x^{j-1}, j = 1, 2, \dots, n+1\right\}, \qquad B_W = \left\{q_i = x^i, i = 1, 2, \dots, n+1\right\}$$

and note that $\dim V = \dim W = n+1$. Let $I\colon V \to W$ be given by

$$I(p)(x) = \int p(x)\mathrm{d}x.$$

The above choice of codomain ensures that the anti-derivative $I(p)$ will result in a polynomial with zero constant term. The matrix representation $A \in \mathbb{R}^{(n+1)\times(n+1)}$ satisfies for all $j = 1, \dots, n+1$

$$I(q_j) = \sum_{i=1}^{n+1} a_{ij}q_i \iff \frac{1}{j}x^j = \sum_{i=1}^{n+1} a_{ij}x^i \iff a_{ij} = \begin{cases} 0 & i \neq j, \\ 1/j & i = j, \end{cases} \iff a_{ij} = \frac{1}{j}\delta_{ij}.$$

The resulting matrix is diagonal and is included below:

$$A = \begin{bmatrix} 1/1 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & 1/2 & 0 & \cdots & \cdots & \vdots \\ \vdots & 0 & 1/3 & 0 & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & 1/n+1 \end{bmatrix} \in \mathbb{R}^{(n+1)\times(n+1)}.$$

10.3 Commutative diagrams

Proposition 10.1 indicates that the action of f on a vector $\mathbf{v}$ is associated with a matrix-vector product $\mathbf{y} = A\mathbf{x}$ involving the coordinate vectors $\mathbf{x}, \mathbf{y}$ of $\mathbf{v}$ and $f(\mathbf{v})$, respectively. We can make this more precise by introducing the concept of **coordinate map** (see Lecture 6). We define the following linear bijections $\varphi_V : V \to \mathbb{R}^n, \varphi_W : W \to \mathbb{R}^m$, via

$$\varphi_V(\mathbf{v}) = \mathbf{x}, \qquad \varphi_W(\mathbf{w}) = \mathbf{y}.$$

We also define the map $\alpha : \mathbb{R}^n \to \mathbb{R}^n$ via $\alpha(\mathbf{x}) = A\mathbf{x}$. Then

$$\varphi_W(f(\mathbf{v})) = \mathbf{y} = A\mathbf{x} = \alpha(\varphi_V(\mathbf{v})) \iff \alpha \circ \varphi_V = \varphi_W \circ f.$$

This is represented in the diagram below, which is known as a **commutative diagram**:

$$\begin{array}{ccc} \mathbf{v} & \xrightarrow{\quad f \quad} & \mathbf{w} = f(\mathbf{v}) \\ {\scriptstyle \varphi_V}\big\downarrow & & \big\downarrow{\scriptstyle \varphi_W} \\ \mathbf{x} & \xrightarrow{\quad \alpha \quad} & \mathbf{y} = A\mathbf{x} \end{array}$$

The term commutative can be best understood when $W = V$: in this case applying the action of f in V followed by the coordinate map φ_V yields the same result as applying the coordinate map followed by the action of f (i.e., of A) in $\mathbb{R}^n$. More generally, following the two paths from $\mathbf{v}$ indicated by the arrows yields the same result: $\mathbf{y}$. Should the maps be invertible (see later), we would be able to draw the arrows also in the opposite directions, thus allowing us to establish one-to-one correspondences. For this reason, commutative diagrams are useful, as they enable to identify more complex (compositions) of actions of linear maps. We will be using this device later when we discuss change of basis.