LECTURE 12

Change of basis

Different bases lead to different representations of vectors as well as matrices associated with a linear map. In this lecture, we examine each case in detail.

12.1 Change of coordinates

Let $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$, $B' = \{\mathbf{v}'_1, \mathbf{v}'_2, \ldots, \mathbf{v}'_n\}$ denote bases for a vector space V. Then any vector $\mathbf{v} \in V$ can be represented in each of the two bases, with different corresponding coordinates:

$$\mathbf{v} = \sum_{j=1}^{n} x_j \mathbf{v}_j = \sum_{i=1}^{n} x'_i \mathbf{v}'_i.$$

It is useful to establish the connection between the two sets of coordinates; we can do this by first considering the connection between the elements of the two bases. To this end, we note that each basis element in B can be written in the basis B' as follows:

$$\mathbf{v}_j = \sum_{i=1}^{n} m_{ij} \mathbf{v}'_i.$$

Using the above representations of $\mathbf{v}$, we find

$$\mathbf{v} = \sum_{j=1}^{n} x_j \mathbf{v}_j = \sum_{j=1}^{n} x_j \sum_{i=1}^{n} m_{ij} \mathbf{v}'_i = \sum_{i=1}^{n} \sum_{j=1}^{n} x_j m_{ij} \mathbf{v}'_i = \sum_{i=1}^{n} x'_i \mathbf{v}'_i$$

so that

$$x'_i = \sum_{i=1}^{n} m_{ij} x_j \Longleftrightarrow \mathbf{x}' = M\mathbf{x}.$$

The matrix M is called the **transition matrix**, as it allows for the change of coordinates from basis B to basis B'. Note that this equivalence can be seen as a special case of the matrix representation of the identity map with respect to different bases, as indicated graphically below (see first remark in L.12):

$$\begin{array}{ccc} \mathbf{v} & \xrightarrow{\;id\;} & \mathbf{v} = id(\mathbf{v}) \\ \varphi \big\downarrow & & \big\downarrow \varphi' \\ \mathbf{x} & \xrightarrow{\;M\;} & \mathbf{x}' = M\mathbf{x} \end{array}$$

Note that $M \neq I_n$ unless $B' = B$, which corresponds to no change of basis or coordinates. Note also that in some situations we may have to be precise with regard to the notation for the transition matrix. Where it is not clear from the context, we will use notation such as $M_{BB'}$ or $\mathbf{x}_{B'} = M_{BB'}\mathbf{x}_B$.

Proposition 12.1 The transition matrix $M_{BB'}$ is invertible.

Proof. Consider the transition matrices from B to B' and from B' to B. We have

$$\mathbf{x}_{B'} = M_{BB'}\mathbf{x}_B, \qquad \mathbf{x}_B = M_{B'B}\mathbf{x}_{B'}.$$

Hence,

$$\mathbf{x}_B = M_{B'B}M_{BB'}\mathbf{x}_B, \qquad \mathbf{x}_{B'} = M_{BB'}M_{B'B}\mathbf{x}_{B'},$$

so that

$$M_{B'B}M_{BB'} = I = M_{BB'}M_{B'B}.$$

By definition, $M_{BB'}$ is invertible with inverse $M_{BB'}^{-1} = M_{B'B}$. ■

Let us consider an example.

Example 12.1 Let

$$B = \left\{1, x, x^2, x^3\right\} =: \{p_i : i = 1, \dots, 4\}, \qquad B' = \left\{1, x, x^2 - \frac{1}{3}, x^3 - \frac{3}{5}x\right\} =: \{q_i, i = 1, \dots, 4\}.$$

The transition matrix $M = M_{BB'} \in \mathbb{R}^{4\times 4}$ has entries m_{ij} given by

$$p_j = \sum_{i=1}^{4} m_{ij} q_i$$

We have

$$\begin{cases} p_1(x) = 1 = m_{11} + m_{21}x + m_{31}\left(x^2 - \frac{1}{3}\right) + m_{41}\left(x^3 - \frac{3}{5}x\right) \\ p_2(x) = x = m_{12} + m_{22}x + m_{32}\left(x^2 - \frac{1}{3}\right) + m_{42}\left(x^3 - \frac{3}{5}x\right) \\ p_3(x) = x^2 = m_{13} + m_{23}x + m_{33}\left(x^2 - \frac{1}{3}\right) + m_{43}\left(x^3 - \frac{3}{5}x\right) \\ p_4(x) = x^3 = m_{14} + m_{24}x + m_{34}\left(x^2 - \frac{1}{3}\right) + m_{44}\left(x^3 - \frac{3}{5}x\right) \end{cases} \Longrightarrow \begin{cases} m_{11} = 1, m_{21} = m_{31} = m_{41} = 0 \\ m_{22} = 1, m_{12} = m_{32} = m_{42} = 0 \\ m_{33} = 1, m_{13} - \frac{1}{3}m_{33} = 0, m_{23} = m_{43} = 0 \\ m_{44} = 1, m_{24} - \frac{3}{5}m_{44} = 0, m_{14} = m_{34} = 0 \end{cases}$$

so that

$$M_{BB'} = \begin{bmatrix} 1 & 0 & 1/3 & 0 \\ 0 & 1 & 0 & 3/5 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Note that an alternative is to compute the entries in the matrix $M_{B'B}$, which is the inverse of $M_{BB'}$:

$$\begin{cases} q_1(x) = 1 & = m_{11} + m_{21}x + m_{31}x^2 + m_{41}x^3 \\ q_2(x) = x & = m_{12} + m_{22}x + m_{32}x^2 + m_{42}x^3 \\ q_3(x) = x^2 - \frac{1}{3} & = m_{13} + m_{23}x + m_{33}x^2 + m_{43}x^3 \\ q_4(x) = x^3 - \frac{3}{5}x & = m_{14} + m_{24}x + m_{34}x^2 + m_{44}x^3 \end{cases} \implies M_{B'B} = \begin{bmatrix} 1 & 0 & -1/3 & 0 \\ 0 & 1 & 0 & -3/5 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

so that

$$M_{BB'} = \begin{bmatrix} 1 & 0 & -1/3 & 0 \\ 0 & 1 & 0 & -3/5 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 & 1/3 & 0 \\ 0 & 1 & 0 & 3/5 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Motivation: the polynomials in the basis B' are Legendre polynomials, which are orthogonal in the inner product

$$\langle p, q \rangle = \int_{-1}^{1} p(x)q(x)\mathrm{d}x.$$

For this reason, they are preferred in many applications and therefore transition matrices need to be computed.

12.2 Change of matrix representation

Let us consider now the matrix representation of a linear map $f : V \mapsto W$ under changes of bases. We will investigate this using diagrams. We first recall the diagram for the matrix representation relative to bases B_V, B_W using the notation indicated above:

$$\begin{tikzcd} \mathbf{v} \arrow[r, "f"] \arrow[d, "\varphi_V"'] & \mathbf{w} = f(\mathbf{v}) \arrow[d, "\varphi_W"] \\ \mathbf{x}_V \arrow[r, "A_{VW}"] & \mathbf{y}_W = A_{VW}\mathbf{x}_V \end{tikzcd}$$

Assume now that we change the basis B_W to B'_W. The resulting diagram is a combination of the previous two diagrams:

$$\begin{tikzcd} \mathbf{v} \arrow[r, "f"] \arrow[d, "\varphi_V"'] & \mathbf{w} = f(\mathbf{v}) \arrow[r, "id_{\mathrm{W}}"] \arrow[d, "\varphi_W"] & \mathbf{w} = id_W(\mathbf{w}) \arrow[d, "\varphi'_W"] \\ \mathbf{x}_V \arrow[r, "A_{VW}"] & \mathbf{y}_W = A_{VW}\mathbf{x}_V \arrow[r, "M_{WW'}"] & \mathbf{y}_{W'} = M_{WW'}\mathbf{y}_W \end{tikzcd}$$

Hence

$$\mathbf{y}_{W'} = M_{WW'}A_{VW}\mathbf{x}_V,$$

so that, by definition, the matrix representation of f with respect to bases B_V, B'_W is

$$A_{VW'} = M_{WW'}A_{VW}.$$

We can derive a similar relation for the case where we change the basis B_V. First, consider the corresponding diagram:

$$\begin{tikzcd} \mathbf{v} \arrow[r, "id_V"] \arrow[d, "\varphi_{V'}"'] & \mathbf{v} = id(\mathbf{v}) \arrow[r, "f"] \arrow[d, "\varphi_V"] & \mathbf{w} = f(\mathbf{v}) \arrow[d, "\varphi_W"] \\ \mathbf{x}_{V'} \arrow[r, "M_{V'V}"] & \mathbf{x}_V = M_{V'V}\mathbf{x}_{V'} \arrow[r, "A_{VW}"] & \mathbf{y}_W = A_{VW}\mathbf{x}_V \end{tikzcd}$$

Thus, $\mathbf{y}_W = A_{VW}M_{V'V}\mathbf{x}_{V'}$ and since $\mathbf{w} = (id_V \circ f)(\mathbf{v}) = f(\mathbf{v})$, we conclude that the matrix representations satisfy $A_{V'W} = A_{VW}M_{V'V}$. The final diagram, corresponding to changing both bases, is included below.

$$\begin{tikzcd}
\mathbf{v} \arrow[r, "id_V"] \arrow[d, "\varphi_{V'}"'] & \mathbf{v} = id(\mathbf{v}) \arrow[r, "f"] \arrow[d, "\varphi_V"] & \mathbf{w} = f(\mathbf{v}) \arrow[r, "id_W"] \arrow[d, "\varphi_W"] & \mathbf{w} = id_W(\mathbf{w}) \arrow[d, "\varphi'_W"] \\
\mathbf{x}_{V'} \arrow[r, "M_{V'V}"] & \mathbf{x}_V = M_{V'V}\mathbf{x}_{V'} \arrow[r, "A_{VW}"] & \mathbf{y}_W = A_{VW}\mathbf{x}_V \arrow[r, "M_{WW'}"] & \mathbf{y}_{W'} = M_{WW'}\mathbf{y}_W
\end{tikzcd}$$

We obtain $\mathbf{y}_{W'} = M_{WW'}A_{VW}M_{V'V}\mathbf{x}_{V'}$, so that

$$A_{V'W'} = M_{WW'}A_{VW}M_{V'V}.$$

Note that by swapping V with V' and W with W' we obtain

$$A_{VW} = M_{W'W}A_{V'W'}M_{VV'}.$$

This relation can also be obtained by using the invertibility of $M_{VV'}$ and $M_{WW'}$. This discussion is summarised in the following theorem.

Theorem 12.2 Let $f : V \mapsto W$ be a linear map between finite-dimensional vector spaces V and W. Let

- B_V, B'_V be bases for V, with corresponding transition matrix $M_{VV'}$;
- B_W, B'_W bases for W, with corresponding transition matrix $M_{WW'}$.

Then the matrix representations of f relative to all four basis combinations satisfy

$$A_{VW} = A_{V'W}M_{VV'} = M^{-1}_{WW'}A_{VW'} = M^{-1}_{WW'}A_{V'W'}M_{VV'}.$$

R If in the above theorem we let $f = id\colon V \to V$, then assuming $A_{VV} = I_n$, we find

$$I_n = M^{-1}_{VV'}A_{V'V'}M_{VV'} \Longrightarrow A_{V'V'} = M_{VV'}I_nM^{-1}_{VV'} = I_n.$$

This confirms that $id : V \to V$ has I_n as matrix representation, irrespective of choice of basis.

Exercise 12.1 Let V, W be vector spaces with dimensions n, m, respectively. Show that the zero map $o : V \to W$ has matrix representation $O_{m,n} \in \mathbb{R}^{m\times n}$, irrespective of choice of basis.

The expression in Theorem 12.2 connecting A_{VW} and $A_{V'W'}$ suggests the following definition:

Definition 12.1 — Matrix equivalence. We say $A, B \in \mathbb{F}^{m\times n}$ are equivalent if $\exists M \in \mathbb{F}^{m\times m}, N \in \mathbb{F}^{n\times n}$, both invertible, such that

$$B = M^{-1}AN.$$

By Theorem 12.2, two matrices are equivalent if they are the matrix representations of the same linear map. [The converse also holds: two equivalent matrices are the representations of some linear map relative to some suitably chosen bases]. The consequence is that any properties inherited from the linear map will be shared by the matrix representations. The first one to note is the rank.

Proposition 12.3 Let $A, B \in \mathbb{F}^{m\times n}$ be equivalent. Then rank $A =$ rank B.

By the rank-nullity formula, we also have that nullity $A =$ nullity B. Another feature that both A and B share is discussed next.

12.3 Canonical forms

Recall that an **equivalence relation**, denoted by $\sim$ is a binary relation on a set S satisfying the following three properties:

i. symmetry: $\alpha \sim \alpha$ for all $\alpha \in S$;
ii reflexivity: if $\alpha \sim \beta$, then $\beta \sim \alpha$ for all $\alpha, \beta \in S$;

iii. transitivity: if $\alpha \sim \beta$ and $\beta \sim \gamma$, then $\alpha \sim \gamma$, for all $\alpha, \beta, \gamma \in S$.

Given an equivalence class, it is natural to ask if there is a representative element in this class. This could be identified through some special (simple) form, or properties. Ideally, the criteria for designating this special form should identify a unique element from the equivalence class. We will refer to this as a **canonical** or **normal** form. We consider this for the case of matrix equivalence.

Proposition 12.4 Matrix equivalence is an equivalence relation on $\mathbb{F}^{m\times n}$.

Proof. Exercise. ■

We conclude that the matrix representations of a linear map belong to the same equivalence class. A candidate for a normal form is a diagonal matrix, due to its simplicity.

Proposition 12.5 Let V, W be vector spaces with dimensions n and m, respectively. Let $f \in \mathcal{L}(V, W)$ have rank r. Then there are bases B_V and B_W relative to which f has matrix representation

$$A_{VW} = \begin{bmatrix} I_r & O_{r,n-r} \\ O_{m-r,r} & O_{m-r,n-r} \end{bmatrix}.$$

Proof. The proof uses the results and notation included in the proof of Theorem 9.8 (the rank-nullity formula). Let $\{B_1, B_2\}$ be a basis for V, where

- $B_1 = \{\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_{n-r}\}$ is a basis for $\ker f$;
- $B_2 = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ is linearly independent.

By the proof of Theorem 9.8, the set $\tilde{B}_1 = \{\mathbf{w}_1 := f(\mathbf{v}_1), \mathbf{w}_2 := f(\mathbf{v}_2), \cdots, \mathbf{w}_r := f(\mathbf{v}_r)\}$ is a basis for $\operatorname{im} f$. Moreover, by Proposition 4.3, $\tilde{B}_1$ is contained in some basis B_W for W, say, $B_W = \left\{\tilde{B}_1, \tilde{B}_2\right\}$ for some set $\tilde{B}_2$. Define $\mathbf{v}_{i+r} := \mathbf{z}_i$ for $i = 1, 2, \ldots, n-r$. Let

$$B_V := \{B_2, B_1\} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}.$$

With this choice of bases for V and W, we find

$$f(\mathbf{v}_i) = \begin{cases} \mathbf{w}_i & 1 \le i \le r, \\ \mathbf{0}_W & r+1 \le i \le n, \end{cases}$$

and the result follows. ■

We can explicitly see that the matrix A_{VW} has rank r, since it only have r linearly independent columns. Thus, the above result indicates that any matrix representation of f, has rank r, where $r = \operatorname{rank} f$.

Definition 12.2 The **canonical form** of the matrix representation of a linear map with rank r is the matrix

$$A^{\text{can}} = \begin{bmatrix} I_r & \\ & \end{bmatrix}.$$