2.6 Matrix Representations of Linear Maps

We have seen that every linear map $F^n \to F^m$ is given by left multiplication by an invertible $m\times n$ matrix $A \in F^{m\times n}$.

Suppose V and W are vector spaces w/ bases $B=(b_1,\dots,b_n)$ and $C=(c_1,\dots,c_m)$ and suppose $L:V\to W$ is a linear map. We have isomorphisms $V\cong F^n$ and $W\cong F^m$.

$$\begin{tikzcd}
v \arrow[r, mapsto] & {[v]_B} \\
V \arrow[r, leftrightarrow, "\sim"] \arrow[d, "L"'] & F^n \arrow[d, "A\cdot"] \\
W \arrow[r, leftrightarrow, "\sim"] & F^m \\
w \arrow[r, mapsto] & {[w]_C}
\end{tikzcd}$$

We can complete the diagram by defining the red arrow to be the composition

$$\begin{tikzcd}
e_i \arrow[r, mapsto] & b_i \arrow[r, mapsto] & L(b_i) \arrow[r, mapsto] & {[L(b_i)]_C} \\
F^n \arrow[r, "\sim"] \arrow[rrr, bend right, "A\cdot"'] & V \arrow[r, "L"] & W \arrow[r, "\sim"] & F^m
\end{tikzcd}$$

This map is given by left multiplication by a matrix $A\in F^{m\times n}$.

2.6.1 Definition

The **matrix of a linear map** $L:V\to W$ with respect to the bases $B=(b_1,\dots,b_n)$ and $C=(c_1,\dots,c_m)$ for V and W, respectively, is the $m\times n$ matrix $[L]_B^C$ whose i^{th} column is

$[L(b_i)]_C$, $1 \le i \le n$. I.e.

$$[L]_B^C = \left([L(b_1)]_C \mid [L(b_2)]_C \mid \cdots \mid [L(b_n)]_C\right).$$

We write $[L]_B$ for $[L]_B^B$. $\square$

Note: $\mathrm{id}_V : V \longrightarrow V$, $v \longmapsto v$ $\qquad [\mathrm{id}]_B^C = T_B^C$

2.6.2 Theorem

Let V and W be vector spaces w/ bases $B = (b_1, \dots, b_n)$ and $C = (c_1, \dots, c_m)$.

(a) For each linear map $L : V \longrightarrow W$, the matrix $[L]_B^C \in F^{m\times n}$ the unique matrix satisfying $[L]_B^C [v]_B = [L(v)]_C$ for all $v \in V$. In other words, $[L]_B^C$ is the unique matrix making the diagram

$$\begin{tikzcd}
V \arrow[r, "\sim"] \arrow[d, "L"'] & F^n \arrow[d, "{[L]_B^C}"] \\
W \arrow[r, "\sim"] & F^m
\end{tikzcd}$$

commute.

(b) The function

$$\begin{aligned} \mathrm{Hom}_F(V, W) &\longrightarrow F^{m\times n} \\ L &\longmapsto [L]_B^C \end{aligned}$$

is a vector space isomorphism.

Proof. (a) $[L]_B^C$ has the desired property by definition. Suppose $A \in F^{m\times n}$ satisfies $A[v]_B = [L(v)]_C$ $\forall v \in V$. Then

$$i^{\text{th}}\text{-column of } A = Ae_i = A[b_i]_B = [L(b_i)]_C = i\text{th column of } [L]_B^C.$$

(b) The function $L \mapsto [L]_B^C$ is linear.

Let $L, K \in \mathrm{Hom}(V, W)$. Then

$$\begin{aligned}
[L+K]_B^C &= \left([(L+K)(b_1)]_C \cdots [(L+K)(b_n)]_C \right) \\
&= \left([L(b_1)+K(b_1)]_C \cdots [L(b_n)+K(b_n)]_C \right) \\
&= \left([L(b_1)]_C + [K(b_1)]_C \cdots [L(b_n)]_C + [K(b_n)]_C \right) \\
&= \left([L(b_1)]_C \cdots [L(b_n)]_C \right) + \left([K(b_1)]_C \cdots [K(b_n)]_C \right) \\
&= [L]_B^C + [K]_B^C
\end{aligned}$$

Similarly, $[\alpha L]_B^C = \alpha [L]_B^C$ for any $\alpha \in F$.
An inverse function $F^{m \times n} \longrightarrow \mathrm{Hom}(V,W)$ is given by sending $A \in F^{m \times n}$ to

$$V \xrightarrow{\sim} F^n \xrightarrow{A\cdot} F^m \xrightarrow{\sim} W.$$

$\square$

2.6.3 Corollary If V and W are finite-dimensional vector spaces, then so is $\mathrm{Hom}(V,W)$ and $\dim_F \mathrm{Hom}(V,W) = \dim_F V \cdot \dim_F W$.

Proof.

$$\dim_F \mathrm{Hom}(V,W) = \dim_F F^{m \times n} = n \cdot m = \dim_F V \cdot \dim_F W.$$

$\square$

2.6.4 Proposition Let V be a finite-dimensional vector space w/ bases A, B and let W be a finite-dimensional vector space w/ bases C and D. If $L: V \longrightarrow W$ is a linear map, then

$$T_C^D [L]_A^C = [L]_B^D T_A^B .$$

Proof. One can prove this w/ a "diagram chase".

$$\begin{tikzcd}
F^n \arrow[rrrr, "{[L]^C_A}"] \arrow[rd, "\sim"] \arrow[dddd, "T^B_A"', "G"] & & & & F^m \arrow[dddd, "T^D_C", "G"'] \\
& & G & & \\
& V \arrow[rr, "L"] \arrow[lddd, "\sim"] & & W \arrow[ruu, "\sim"'] \arrow[rdd, "\sim"] & \\
& & G & & \\
F^n \arrow[rrrr, "{[L]^D_B}"'] & & & & F^m
\end{tikzcd}$$

Path 4

Path 5

Path 2

Path 3

Path 1

The squares + triangles commute by 2.6.2 (a)

The claim is that the blue square commutes. To prove this, use the commuting triangles and squares.

Path 1 = Path 2 = Path 3 = Path 4 = Path 5.

Alternate proof:

$$\begin{aligned}
[L]^D_B \, T^B_A \, [v]_A = [L]^D_B [v]_B &= [L(v)]_D \\
&= T^D_C \, [L(v)]_C \\
&= T^D_C \, [L]^C_A [v]_A .
\end{aligned}$$

□

2.6.5 Example For any field F, formal diff. of polynomials defines a linear map $D : F_n[x] \to F_n[x]$

(a) Compute $[D]_E$ where $E = (1, x, \dots, x^n)$.

$$[D]_E = \left([0]_E \ [1]_E \ [2x]_E \cdots\cdots [nx^{n-1}]_E \right)$$

$$= \begin{pmatrix}
0 & 1 & 0 & 0 & \cdots & 0 \\
0 & 0 & 2 & 0 & & 0 \\
0 & 0 & 0 & 3 & & 0 \\
\vdots & \vdots & \vdots & 0 & \ddots & \vdots \\
 & & & \vdots & & n \\
0 & 0 & 0 & 0 & \cdots & 0
\end{pmatrix}$$

(b) Consider $n=2$, $F=\mathbb{Z}_2$, $B=(1+x, x+x^2, x)$, $C=(1, x+x^2, 1+x^2)$. Compute $[D]^C_B$ using Prop 2.6.4: $[D]^C_B = (T^E_C)^{-1}[D]^E_B$.

$$[D]^E_B = \begin{pmatrix} [1]_E & [1]_E & [1]_E \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$(T^E_C)^{-1} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

$$[D]^C_B = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$\square$

2.6.6 Proposition Suppose $L: U \longrightarrow V$ and $K: V \longrightarrow W$ are linear maps between finite-dimensional vector spaces U, V, W with bases B, C, D, respectively. Then

$$[KL]^D_B = [K]^D_C \, [L]^C_B$$

Proof. Diagram chase:

$$\begin{tikzcd} U \arrow[r, "L"] \arrow[d, "S"'] \arrow[dr, phantom, "\circlearrowleft"] & V \arrow[r, "K"] \arrow[d, "S"'] \arrow[dr, phantom, "\circlearrowleft"] & W \arrow[d, "S"'] \\ F^n \arrow[r, "{[L]^C_B}"'] & F^m \arrow[r, "{[K]^D_C}"'] & F^k \end{tikzcd}$$

Both squares commute by 2.6.2(a). Therefore, the outer square

$$\begin{tikzcd} U \arrow[r, "KL"] \arrow[d, "S"'] \arrow[dr, phantom, "\circlearrowleft"] & W \arrow[d] \\ F^n \arrow[r, "{[K]^D_C [L]^C_B}"'] & F^k \end{tikzcd}$$

But by Thm 2.6.2.(a), the unique matrix making the blue square commute is $[KL]_B^D$. Thus,

$$[K]_C^D [L]_B^C = [KL]_B^D .$$

□

2.6.7 Corollary Let V be a finite-dimensional vector space w/ basis B. The function

$$\begin{aligned} \mathrm{End}_F(V) &\longrightarrow F^{n\times n} \\ L &\longmapsto [L]_B \end{aligned}$$

defines an isomorphism of (unital) F-algebras.

Proof. We already know the function is a vector space isomorphism by 2.6.2(b). An algebra isomorphism is a v.s. isomorphism which preserves the product. We just proved this in 2.6.6. It is unital because $[\mathrm{id}_V]_B = I_n$.

□

2.7 Similarity of Linear Operators

Recall: A linear map $L: V \longrightarrow V$ is called a linear operator. The set of all linear operators is denoted $\mathrm{End}_F(V) := \mathrm{Hom}_F(V,V)$. We proved that every linear operator can be represented by a matrix $[L]_B$ by choosing a basis B for V. But different bases will produce different matrices for the same linear operator.

Question: How can we understand the many different matrix representations of the same linear operator L?

Answer: Similarity of matrices.

2.7.1 Definition Two matrices $X, Y \in F^{n\times n}$ are called **similar** if there exists an invertible matrix $P \in F^{n\times n}$ such that $Y = P^{-1}XP$. We write $X \sim Y$ if X and Y are similar.

$\square$

Similarity defines an equivalence relation on $F^{n\times n}$.

2.7.2 Theorem Let V be an F-vector space of dimension n. Two matrices $X, Y \in F^{n\times n}$ are similar if and only if they represent the same linear operator on V (with respect to possibly different bases for V).

Proof. ($\Rightarrow$) Suppose $X \sim Y$. Then there exists an invertible matrix $P \in F^{n\times n}$ s.t. $Y = P^{-1}XP$. We need to find a linear operator $L: V \to V$ and bases B, C for V s.t. $X = [L]_B$ and $Y = [L]_C$.

Let B be any basis for V. Using the isomorphism

$$\begin{tikzcd} \mathrm{End}(V) \arrow[r] & F^{n\times n} \\ L \arrow[r, mapsto] & [L]_B \end{tikzcd}$$

choose $L \in \mathrm{End}(V)$ s.t. $X = [L]_B$. Since P is invertible, we have an isomorphism

$$\begin{tikzcd} \varphi : F^n \arrow[r, "\overset{P\cdot}{\sim}", "e_i \mapsto Pe_i"'] & F^n \arrow[r, "\overset{e_i \mapsto b_i}{\sim}"] & V. \end{tikzcd}$$

Because φ is an isomorphism, $(\varphi(e_1), \ldots, \varphi(e_n)) =: C$ is a basis for V. We claim $[L]_C = Y$. Consider the diagram

$$\begin{tikzcd} F^n \arrow[r, "P\cdot"] \arrow[d, "Y\cdot"'] \arrow[dr, phantom, "\circlearrowleft"] & F^n \arrow[r, "e_i \mapsto b_i"] \arrow[d, "X\cdot"] \arrow[dr, phantom, "\circlearrowleft"] & V \arrow[d, "L"] \\ F^n \arrow[r, "P\cdot"'] & F^n \arrow[r, "e_i \mapsto b_i"'] & V \end{tikzcd}$$

[$XP = PY$
$\iff$
$Y = P^{-1}XP$]

$X = [L]_B$
$+$ Thm 2.6.2.

Thus, the outer square

$$\begin{tikzcd} F^n \arrow[r, "\overset{\varphi}{\sim}"] \arrow[d, "Y\cdot"'] \arrow[dr, phantom, "\circlearrowleft"] & V \arrow[d, "L"] \\ F \arrow[r, "\sim", "\varphi"'] & V \end{tikzcd}$$

commutes. But φ is just the coordinate isomorphism corresponding to the basis C. The unique matrix making the diagram commute is $[L]_C$ by 2.6.2. Thus, $Y = [L]_C$.

($\Leftarrow$) Suppose $X = [L]_B$ and $Y = [L]_C$ for some $L \in \mathrm{End}(V)$ and some bases B, C for V. Then

$$X = [L]_B = T^B_C [L]_C T^C_B$$

$$= (T^C_B)^{-1} [L]_C T^C_B = (T^C_B)^{-1} Y T^C_B . \square$$

2.8 Direct Sums and Quotients

2.8.1 Definition Let $V_1, \dots, V_n$ be vector spaces over F. Their (external) direct sum is the set

$$\bigoplus_{i=1}^{n} V_i := \{ (v_1, v_2, \dots, v_n) \mid v_i \in V_i , i = 1, \dots, n \}$$

together with the operations

(addition) $(v_1, \dots, v_n) + (w_1, \dots, w_n) := (v_1 + w_1 , \dots , w_1 + w_n)$

(scalar mult) $\alpha(v_1, \dots, v_n) := (\alpha v_1, \dots, \alpha v_n)$.

□

Remark: It is easy to check that $\bigoplus_{i=1}^{n} V_i$ is a vector space. As sets, $\bigoplus_{i=1}^{n} V_i$ is just the cartesian product $\prod_{i=1}^{n} V_i$. But we will see soon that there is good reason to use "$\oplus$" instead of "Π".

Remark: One can define direct sum for any family $\{V_i \mid i \in I\}$ of vector spaces over F. One starts by defining the direct product

$$\prod_{i \in I} V_i := \{ f : I \to \bigcup_{i \in I} V_i \mid f(i) \in V_i \ \forall i \in I \} .$$

This is a vector space w/ "coordinate-wise" addition and scalar multiplication

$$(f+g)(i) := f(i)+g(i)$$
$$(\alpha f)(i) := \alpha f(i).$$

The direct sum is the subspace of $\prod_{i\in I} V_i$ consisting of all functions for which $f(i)=0$ for all but finitely many $i\in I$

$$\bigoplus_{i\in I} V_i := \left\{ f\in\prod_{i\in I} V_i \;\middle|\; \begin{array}{l} f(i)=0 \text{ for all but} \\ \text{finitely many } i\in I\end{array}\right\}.$$

Because each function has finitely many nonzero outputs it is customary to write $f\in\bigoplus_{i\in I}V_i$ as a formal sum $\sum_{i\in I} f_i$ $(f_i := f(i))$. Using this notation,

$$\sum_{i\in I} f_i + \sum_{i\in I} g_i := \sum_{i\in I} f_i+g_i$$
$$\alpha\sum_{i\in I} f_i = \sum_{i\in I}\alpha f_i.$$

When I is finite, $\bigoplus_{i\in I}V_i = \prod_{i\in I}V_i$.

2.8.2 Example Let F be a field. Then

$$F^n = \{(a_1,\dots,a_n) \mid a_i\in F\}$$
$$= \underbrace{F\oplus\cdots\oplus F}_{n\text{-summands}}$$

More generally, let X be any set. A long time ago I defined a subspace $F^{(X)}$ of F^X

$$F^{(X)} = \left\{f:X\to F \;\middle|\; \begin{array}{l} f(x)=0 \text{ for all but} \\ \text{finitely many } x\in X\end{array}\right\}.$$

We call $F^{(X)}$ the free vector space w/ basis X. Since $F = \bigcup_{x \in X} F$, we see that $F^X = \prod_{x \in X} F$ and $F^{(X)} = \bigoplus_{x \in X} F$.

□

2.8.3 Proposition The set $\{(0,0,\ldots,v,\ldots,0) \mid v \in V_i\}$ is a subspace of $\bigoplus_{i=1}^{n} V_i$ that is isomorphic to V_i.

Proof. Clearly the set is a subspace. Define $\iota_i : V_i \longrightarrow \bigoplus_{i=1}^{n} V_i$ via $\iota_i(v) = (0,0,\ldots,v,\ldots,0)$. Clearly, ι_i is linear, injective, and

$$\operatorname{im}(\iota_i) = \{(0,0,\ldots,v,\ldots,0) \mid v \in V_i\}.$$

□

Remark: For this reason, we usually identify V_i w/ the subspace $\{(0,0,\ldots,v,\ldots,0) \mid v \in V_i\}$ of $\bigoplus_{i=1}^{n} V_i$. In this way, we can think of V_i as being a genuine subspace of $\bigoplus_{i=1}^{n} V_i$. The subspaces $V_1, \ldots, V_n$ of $\bigoplus_{i=1}^{n} V_i$ are special:

(a) $\bigoplus_{i=1}^{n} V_i = \sum_{i=1}^{n} V_i$

(b) For each i, $V_i \cap \left(\sum_{j \neq i} V_j \right) = \{0\}$.

2.8.4 Definition Let V be a vector space w/ subspaces $V_1, \dots, V_n$. We say that V is the **(internal) direct sum** of $V_1, \dots, V_n$ if

(a) $V = \sum_{i=1}^{n} V_i$

(b) For each i, $V_i \cap \left(\sum_{j \neq i} V_j \right) = \{0\}$

□

2.8.5 Proposition Let V be a vector space and $V_1, \dots, V_n$ subspaces. The following are equivalent.

(a) V is the internal direct sum of $V_1, \dots, V_n$

(b) For each $v \in V$, there are unique vectors $v_1, \dots, v_n$ w/ $v_i \in V_i$, $i = 1, \dots, n$ such that $v = \sum_{i=1}^{n} v_i$.

(c) There is an isomorphism $\varphi : V \longrightarrow \bigoplus_{i=1}^{n} V_i$ (from V to the external direct sum) sending each subspace $V_i \subseteq V$ to $\{(0, \dots, v, \dots, 0) \mid v \in V\} \subseteq \bigoplus_{i=1} V_i$.

Proof. ((c) ⇒ (a)) Follows from remarks. $\sum v_i - \sum w_i = v - v = 0$

((a) ⇒ (b)) let $v \in V$. Since $V = \sum_{i=1}^{n} V_i$, $v = \sum_{i=1}^{n} v_i$ where $v_i \in V$. Suppose $v = \sum_{i=1}^{n} w_i$ where $w_i \in V_i$.

The claim is that $v_i = w_i$ for all $i = 1, \dots, n$. Then

$$V_i \ni v_i - w_i = \sum_{j \neq i} w_i - \sum_{j \neq i} v_i = \sum_{j \neq i} w_i - v_i \in \sum_{j \neq i} V_j$$

Thus, $v_i - w_i \in V_i \cap \left(\sum_{j \neq i} V_j \right) = \{0\}$. this shows that $v_i - w_i = 0$ for all $i = 1, \dots, n$.

((b) ⇒ (c)) Assume (b). Define $f : V \longrightarrow \bigoplus_{i=1}^{n} V_i$ as follows. Given $v \in V$, there are unique vectors $v_1 \in V_1, \dots, v_n \in V_n$ such that $v = \sum_{i=1}^{n} v_i$. So define $f(v) = (v_1, \dots, v_n)$. By uniqueness, it is a function. Clearly, f is linear. An inverse $\psi : \bigoplus_{i=1}^{n} V_i \to V$ given by $\psi((v_1, \dots, v_n)) = \sum_{i=1}^{n} v_i$. Further,

$v_i \in V_i$

$$f(V_i) = \{(0, \dots, v, \dots, 0) \mid v \in V_i\}.$$

□

Notation: Because of Prop 2.8.4., we write

$$V = \bigoplus_{i=1}^{n} V_i$$

whenever V is the internal direct sum of subspaces $V_1, \dots, V_n$.