8 Dimension Formula

8.1 Review

Last time, we ended off with the definition of linear transformations.

8.2 Matrix of Linear Transformations

Given a linear transformation $T : V \to W$, we know

$$T(a_1 \vec{v}_1 + \cdots + a_n \vec{v}_n) = a_1 T(\vec{v}_1) + \cdots + a_n T(\vec{v}_n).$$

Then, given any basis $\vec{v}_1, \ldots, \vec{v}_n$ for V, the above property tell us that T is completely determined by the values of $T(\vec{v}_1), \ldots, T(\vec{v}_n)$.

Example 8.1

Last time, we discussed a linear transformation from column vectors to another vector space. In particular, if we have a basis $\vec{w}_1, \ldots, \vec{w}_m$ of W, then we can create the following linear transformation:

$$B : F^n \to W$$
$$\vec{e}_i \mapsto \vec{w}_i$$
$$(a_1, \cdots, a_n) \mapsto \sum a_i \vec{w}_i,$$

where $\vec{e}_i$ is the zero column vector with a 1 in the ith position. This is an isomorphism due to choosing $\vec{w}_i$ to be a basis. The inverse map $B^{-1}(\vec{w}) = (a_1, \cdots, a_n)$ sends a vector to the *coordinates* of $\vec{w}$ for the given basis.

Example 8.2

There is a bijection between matrices $A \in \mathrm{Mat}_{m \times n}(F)$ and linear transformations $T : F^n \to F^m$. For every matrix A, it can be mapped to the linear transformation $T = A \cdot \vec{x}$. For every transformation T, it can be mapped to

$$A = \begin{pmatrix} T(\vec{e}_1) & T(\vec{e}_2) & \cdots & T(\vec{e}_n) \end{pmatrix}.$$

Both the set of matrices and the set of linear transformations are vector spaces, so this defines an isomorphism between two vector spaces. We will switch between these notions frequently.

As a special case, if we have an isomorphism $T : F^n \xrightarrow{\sim} F^m$, then this forces $m = n$ and the corresponding matrix A must be in $\mathrm{GL}_n(F)$.

From Example 8.1, suppose we have two different bases of V and create their corresponding linear transformations B and B'. Let the corresponding bases be $\{\vec{v}_1, \ldots, \vec{v}_n\}$ and $\{\vec{w}_1, \ldots, \vec{w}_n\}$ respectively. Then, $P := B^{-1} \circ B'$ defines a mapping from F^n to F^n, and we must have $P \in GL_n(F)$. Furthermore, $B' = B \circ P$. These relations can be seen by following the arrows in the below diagram:

$$F^n \xrightarrow{\ B\ } V$$
$$\underset{P}{\swarrow} \quad \underset{B'}{\uparrow}$$
$$F^n$$

We can figure out the columns of P as well:

$$P(\vec{e}_i) = B^{-1}(B'(\vec{e}_i)) = B^{-1}(\vec{w}_i).$$

We know $B^{-1}(\vec{w}_i)$ is just the coordinates of $\vec{w}_i$ for the basis $\{\vec{v}_1, \ldots, \vec{v}_n\}$. Similarly, we can figure out the columns of P^{-1} by taking the each $\vec{v}_i$ and writing it in terms of the basis of $\vec{w}$'s.

Given a vector $\vec{v} \in V$, we can write the coordinates $\vec{x} = B^{-1}(\vec{v})$ and $\vec{x}' = (B')^{-1}(\vec{v})$. The coordinates are related by $P \vec{x}' = \vec{x}$ and $P^{-1} \vec{x} = \vec{x}'$ by using our expression for P.

For any finite-dimensional vector space, by picking a basis, we can write every vector in terms of coordinates. Then, for every transformation $T : V \to W$, we can write it as a matrix A using coordinates. Suppose we have the coordinate maps $B : F^n \to V$ and $C : F^m \to W$ and the respective bases are $\{\vec{v}_i\}$ and $\{\vec{w}_i\}$.

$$
\begin{array}{ccc}
V & \xrightarrow{\ T\ } & W \\
{\scriptstyle B}\uparrow & & {\scriptstyle C}\uparrow \\
F^n & \xrightarrow{\ A\ } & F^m
\end{array}
$$

Following the arrows, we have $A := C^{-1} \circ T \circ B \in \mathrm{Mat}_{m \times n}(F)$. To write down A, we can find the ith column of A by writing $T(\vec{v}_i)$ in terms of the $\vec{w}$'s.

Example 8.3

Let's see an example of computing A. Let V be the set of complex functions such that $f''(t) = f(t)$. Let W be the set of complex functions such that $f''(t) = -f(t)$. We can define the transformation

$$
T : V \to W
$$
$$
f(t) \mapsto f(it).
$$

Note that T looks nothing like a matrix right now, but we can turn it into one by picking bases for V and W. One choice of basis is $V = \mathrm{Span}(e^t, e^{-t})$ and $W = \mathrm{Span}(\cos t, \sin t)$.

Then, by writing $T(e^t) = e^{it} = \cos t + i \sin t$ and $T(e^{-t}) = e^{-it} = \cos t + -i \sin t$, we have found the columns of A:

$$
A = \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}.
$$

If we had chosen a different basis for $W = \mathrm{Span}(e^{it}, e^{-it})$, then our matrix A' would just be the identity matrix.

Guiding Question

We have seen that the same linear transformation leads to different A, so can we pick bases so that A "looks very nice"? For example, in the previous example, we were able to make A the identity matrix by picking a different basis.

We will answer this question at the end of the next section.

8.3 Dimension Formula

Note that linear transformations are very similar to group homomorphisms. They are both mappings that preserve the structure that we care about. We can define and prove similar results as the ones we showed for group homomorphisms.

Definition 8.4

Given a linear transformation $T : V \to W$, we can define the **kernel** and **image**.

$$
\ker(T) := \{\vec{v} \mid T(\vec{v}) = \vec{0}\}
$$
$$
\mathrm{im}(T) := \{\vec{w} \mid T(\vec{v})\vec{w} \text{ for some } \vec{v} \in V\}
$$

By similar logic to group homomorphisms, these are vector *subspaces* of V and W respectively. We also define the **nullity** and the **rank** as the dimension of the kernel and image respectively.

Theorem 8.5 (Dimension formula)

Given $T : V \to W$,
$$
\dim(\ker T) + \dim(\mathrm{im}\, T) = \dim(V).
$$

This is somewhat reminiscent of the theorem on groups $|G| = |\ker(G)||\mathrm{im}(G)|$.

Proof. Pick $\vec{v}_1, \ldots, \vec{v}_k$ as a basis for $\ker(T)$. From a theorem from last class, we can add vectors $\vec{v}_{k+1}, \ldots, \vec{v}_n$ to get a basis for V where $n = \dim V$ and $k = \dim(\ker(T))$.

Let $\vec{w}_i := T(\vec{v}_i)$. By the definition of kernel, $\vec{w}_i = 0$ for $1 \leq i \leq k$. We will show that $\{\vec{w}_{k+1}, \ldots, \vec{w}_n\}$ is a basis for $\operatorname{im}(T)$, so that
$$\operatorname{rank}(T) = n - k = \dim(V) - \dim(\ker(T)).$$

To prove that it is a basis, we need to show that they are linearly independent and they span $\operatorname{im}(T)$. For the span,

$$\begin{aligned}
\operatorname{im}(T) &= \operatorname{Span}(T(\vec{v}_1), \ldots, T(\vec{v}_n)) \\
&= \operatorname{Span}(T(\vec{v}_{k+1}), \ldots, T(\vec{v}_n)) \\
&= \operatorname{Span}(\vec{w}_{k+1}, \ldots, \vec{w}_n).
\end{aligned}$$

For linear independence, we consider the solutions to:
$$a_{k+1}\vec{w}_{k+1} + \cdots + a_n\vec{w}_n = \vec{0}.$$

This implies that
$$T(a_{k+1}\vec{v}_{k+1} + \cdots + a_n\vec{v}_n) = \vec{0}$$

and thus
$$a_{k+1}\vec{v}_{k+1} + \cdots + a_n\vec{v}_n \in \ker(T).$$

Since $\vec{v}_1, \ldots, \vec{v}_k$ is a basis for the kernel, there must exist coefficients $a_1, \ldots, a_k$ such that
$$a_{k+1}\vec{v}_{k+1} + \cdots + a_n\vec{v}_n = a_1\vec{v}_1 + \cdots + a_k\vec{v}_k.$$

However, this forces $a_i = 0$ for all i since the $\vec{v}$'s form a basis for V. Therefore, $\{\vec{w}_{k+1}, \ldots, \vec{w}_n\}$ are linearly independent and thus a basis of $\operatorname{im}(T)$. $\qquad\square$

The proof of the dimension formula shows a bit more. Using the same notation as in the proof, take a basis for V to be $\vec{v}_{k+1}, \ldots, \vec{v}_n, \vec{v}_1, \ldots, \vec{v}_k$. This is essentially the same basis, but permuted so that the coordinate vectors are also permuted. We extend the basis for $\operatorname{im}(T)$ to a basis for W with the vectors $\vec{w}_{k+1}, \ldots, \vec{w}_n, \vec{u}_1, \ldots, \vec{u}_r$. Then we find the matrix for T by writing down the coordinates of $T(\vec{v}_i)$ with respect to the $\vec{w}$'s. When $k + 1 \leq i \leq n$, $T(\vec{v}_i) = \vec{w}_i$. When $1 \leq i \leq k$, $T(\vec{v}_i) = 0$. Our matrix for T is particularly simple (written in block form):
$$A = \begin{pmatrix} I_{n-k} & 0 \\ 0 & 0 \end{pmatrix}.$$

Corollary 8.6

For any linear transformation, we can write its matrix in the above form for some choice of basis for V and W.

Corollary 8.7

As a special case, if we already are given a matrix $M \in \operatorname{Mat}_{m \times n}$ representing a linear transformation from $F^n \to F^m$, then there exists change of basis matrices $P \in \operatorname{GL}_n(F), Q \in \operatorname{GL}_m(F)$ such that $Q^{-1}MP$ is in the above form. Pictorially, this looks like:

$$\begin{array}{ccc}
F^n & \xrightarrow{\;M\;} & F^m \\
{\scriptstyle P}\Big\uparrow & & {\scriptstyle Q}\Big\uparrow \\
F^n & \xrightarrow{\;A\;} & F^m
\end{array}$$