Lecture 4 - More Linear Transformations

Saw that to give a basis $\mathcal{B}$ for V is to give an isomorphism

$$\mathbb{F}\mathcal{B} \to V$$

specified by taking elements of $\mathcal{B}$ to the corresponding element of $\mathcal{B}$. While this does give us a simpler schema, we need a small amount extra to bring in vectors and matrices.

Def An <u>ordered basis</u> is a basis equipped with a linear order: if $u, v \in \mathcal{B}$, then $u < v$, $u > v$ or $u = v$.

We have a standard linearly ordered finite set of any cardinality:

$$\{1, \dots, n\} \text{ with the standard } <.$$

Given any finite linearly ordered set X of cardinality n, we get an ordered bijection $X \longleftrightarrow \{1, \dots, n\}$ by

$$\begin{array}{ccc} \text{smallest element} & \longleftrightarrow & 1 \\ \vdots & & \vdots \\ \text{biggest element} & \longleftrightarrow & n \end{array}$$

Since $\{1, \dots, n\}$ is ordered, the direct sum over it has a prefered order for writing the basis elements in the sums:

$$\bigoplus_{\{1,\dots,n\}} \mathbb{F} = \mathbb{F}^n = \{(a_1, \dots, a_n)^t \mid a_i \in \mathbb{F}\}.$$

Thus the elements of $\mathbb{F}^n$ are column vectors. Following this chain back, we see the following proposition.

Prop If $\mathcal{B}$ is an ordered basis of V of cardinality n, then $\mathcal{B}$ gives an isomorphism

$$\mathbb{F}^n \xrightarrow{\phi_{\mathcal{B}}} V.$$

This lets us understand vector spaces in an even simpler way. It's useful, though, because it plays nicely with linear transformations.

Prop A linear transformation $\mathbb{F}^n \to \mathbb{F}^m$ is given by an $m \times n$ matrix, and the action is matrix multiplication

Pf: Let L be a linear transformation $\mathbb{F}^n \to \mathbb{F}^m$. We specify L by giving its values on the standard basis vectors

$$\bar{e}_i = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} i. \qquad L(\bar{e}_i) \in \mathbb{F}^m \text{ is some vector.}$$

Then
$$L(\bar{x}) = \left[L(\bar{e}_1) \,\vdots\, L(\bar{e}_2) \,\vdots\, \cdots \,\vdots\, L(\bar{e}_n) \right] \cdot \bar{x}, \quad \text{since}$$

$$\bar{x} = \sum a_i \bar{e}_i = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix},$$ and matrix multiplication $A \cdot \bar{v}$ forms the linear combination of the columns of A w/ coefs from $\bar{x}$. $\boxtimes$

Thus for $\mathbb{F}^n \to \mathbb{F}^m$, linear transformations are the same thing as matrices. Giving an ordered basis gives an identification w/ coordinate vectors, so we get matrices for more general linear transformations.

Prop Let V have ordered basis $\mathcal{B}$ and W have ordered basis $\mathcal{C}$, and let $L: V \to W$ be a linear transformation. Then there is a matrix ${}_{\mathcal{C}}[L]_{\mathcal{B}}$ s.t.

$$\begin{array}{ccc} V & \xrightarrow{L} & W \\ \phi_{\mathcal{B}} \uparrow & & \uparrow \phi_{\mathcal{C}} \\ \mathbb{F}^n & \xrightarrow{{}_{\mathcal{C}}[L]_{\mathcal{B}}} & \mathbb{F}^m \end{array} .$$

The last part means that for any $\bar{v} \in \mathbb{F}^n$,

$$L(\phi_B(\bar{v})) = \phi_e({}_e[L]_B \cdot \bar{v}).$$ Since ϕ is an isomorphism, this really says L is the same data as the matrix.

The way we produce matrices from linear transformations for $\mathbb{F}^n \to \mathbb{F}^m$ tells us how to show this:

Write $$L(\bar{b}_i) = a_{1i}\bar{c}_1 + a_{2i}\bar{c}_2 + \dots + a_{mi}\bar{c}_m,$$

and then recognize that the right-hand side is $\phi_e\left(\begin{bmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{bmatrix}\right)$ while the left-hand side is $L(\phi_B(\bar{e}_i))$

So $${}_e[L]_B = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix},$$ and the ith column is the coefficient vector for $L(\bar{b}_i)$.

__Remark__ Since ϕ is an isomorphism, it is invertible. Then

$$\phi_e^{-1} L \phi_B : \mathbb{F}^n \longrightarrow \mathbb{F}^m$$ is a linear transformation and hence a matrix. This is ${}_e[L]_B$.

To prevent confusion, it is helpful to label the columns with the elements of $\mathcal{B}$ and the rows with elements of $\mathcal{C}$.

This is the start of a very pretty story.

__Def__ Let $L(U,V)$ be the set of all linear transformations from $U \to V$.

__Prop__ With the operations $(a\cdot S + b\cdot T)(\bar{u}) = a\cdot S(\bar{u}) + b\cdot T(\bar{u})$, $L(U,V)$ is a vector space.

Pf: Homework.

__Remark__ It is actually a subspace of the vector space of all functions $U \to V$ $\left(= \prod_U V\right)$.

This remark makes the previous problem easier. Now a restatement of our earlier proposition.

Thm A choice of ordered basis $\mathcal{B}$ for U and $\mathcal{C}$ for V gives an isomorphism

$$L(U,V) \longrightarrow M_{m\times n}(\mathbb{F})$$
$$L \longmapsto {}_{\mathcal{C}}[L]_{\mathcal{B}}.$$

Pf of linearity: Since matrix operations are coordinatewise, it suffices to check that the i^{th} column of $aS+bT$ is the corresponding linear combination of i^{th} columns of S and T. However, the i^{th} column is the vector of coefficients of $L(\bar{b}_i)$, so it suffices to show that the coefficients of $\bar{c}_j$ in $(aS+bT)(\bar{b}_i)$ and $aS(\bar{b}_i)+bT(\bar{b}_i)$ agree. Stated in this way, it's obvious.

We'll delay showing it's bijective for a time.

Def Let $L: U\to V$ be linear. Then the image of L is the set-theoretic image: $\mathrm{Im}(L)$. The kernel of L is
$$\ker(L)=\{\bar{u}\in U \mid L(\bar{u})=\bar{0}\}.$$

Prop $\mathrm{Im}(L)$ is a subspace of V, $\ker(L)$ is a subspace of U.

Pf The first is essentially obvious. For the second:

$\bar{0}\in\ker(L)$. & if $\bar{u},\bar{v}\in\ker(L)$, then
$$L(a\bar{u}+b\bar{v}) = aL(\bar{u})+bL(\bar{v}) = a\cdot\bar{0}+b\cdot\bar{0}=\bar{0}.$$
$$\Rightarrow a\cdot\bar{u}+b\cdot\bar{v}\in\ker(L). \qquad \square$$

Remark The same proof shows that if W is a subspace of V, then $L^{-1}W=\{\bar{u}\mid L(\bar{u})\in W\}$ is a subspace.

The kernel and image play important roles:

- the kernel measures how far L is from being injective.
- the image measures how close L is to being surjective.

Prop L is a monomorphism if and only if $\ker(L)=\{\bar{0}\}$.

Pf: Assume $L(\bar{u})=L(\bar{v})$. Then since L is linear, we can rewrite this as $L(\bar{u}-\bar{v})=\bar{0}$. So

$$L(\bar{u})=L(\bar{v}) \longleftrightarrow \bar{u}-\bar{v}\in \ker(L).$$ Thus if $\ker(L)=\{\bar{0}\}$, then $L(\bar{u})=L(\bar{v}) \Rightarrow \bar{u}=\bar{v}$; if L is injective, then $L(\bar{u})=\bar{0} \Rightarrow \bar{u}=\bar{0}$, so $\ker(L)=\{\bar{0}\}$. $\boxtimes$

We'll return to the idea that the value of L on $\bar{u}$ is determined only up to summands in the kernel later.

For now, choose a complementary subspace to $\ker(L)$ in U:

$$U = \ker(U) \oplus U'$$

We can easily do this by choosing a basis for $\ker(U)$: $\bar{u}_1,\ldots,\bar{u}_m$ and extending it to a basis of U by adding extra vectors: $\bar{u}_{m+1},\ldots,\bar{u}_n$. Then $U'=\langle \bar{u}_{m+1},\ldots,\bar{u}_n\rangle$.

Prop The Image of L restricted to U' is the image of L. The kernel of L restricted to U' is $\{\bar{0}\}$.

Pf: If $\bar{w}\in \mathrm{Im}(L)$, then can find $\bar{v}\in U$ s.t. $L(\bar{v})=\bar{w}$.

Write $\bar{v}=a_1\bar{u}_1+\ldots+a_m\bar{u}_m+a_{m+1}\bar{u}_{m+1}+\ldots+a_n\bar{u}_n$.

Then
$$\begin{aligned} L(\bar{v}) &= a_1\cancel{L(\bar{u}_1)}+\ldots+a_m\cancel{L(\bar{u}_m)}+L(a_{m+1}\bar{u}_{m+1}+\ldots+a_n\bar{u}_n) \\ &= L(a_{m+1}\bar{u}_{m+1}+\ldots+a_n\bar{u}_n) \\ &\in \mathrm{Im}(L|_{U'}). \end{aligned}$$

Similarly, if $\bar{v}\in \ker(L|_{U'})$, then $\bar{v}\in\ker(L)\cap U'$. The sum is direct, so $\bar{v}=\bar{0}$. $\boxtimes$

<u>Cor</u> (Rank-Nullity Theorem)

$$\dim U = \dim(\ker L) + \dim(\operatorname{Im}(L)).$$